

# Entropic Continuity Bounds & Eventually Entanglement-Breaking Channels



Eric P. Hanson

Department of Applied Mathematics and Theoretical Physics  
University of Cambridge

This dissertation is submitted for the degree of  
*Doctor of Philosophy*



To my family, and to Hang Lu.



## Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University of similar institution except as declared in the Preface and specified in the text.

Eric P. Hanson

April 2020



## Acknowledgements

I would like to sincerely thank Nilanjana Datta and Yan Pautrat for the excellent supervision over the past few years. Nilanjana always made time to discuss things with me and I benefited from many blackboard discussions with her. And while I could not meet Yan as frequently due to the distance to Paris, it was a pleasure whenever we met.

Over the course of my PhD, I have benefited tremendously from collaborations with Koenraad Audenaert, Andreas Bluhm, Hao-Chung Cheng, Nilanjana Datta, Daniel Stilck França, Min-Hsiu Hsieh, Alain Joye, Vishal Katariya, Yan Pautrat, Renaud Raquépas, Cambyse Rouzé, Mark Wilde, and Jiazheng Zhu. I hope to work with you all again. I have also been lucky to have been able to have interesting mathematical discussions with many people, including Simon Andréys, Koji Azuma, Ivan Bardet, Johannes Bausch, Ángela Capel Cuevas, Michele Dall’Arno, Kun Fang, Hamza Fawzi, Lisa Hänggli, Christoph Hirche, Joel Klassen, Robert König, Felix Leditzky, Māris Ozols, Sam Power, and many others. I’d also like to thank Hamza Fawzi and Mark Wilde for serving as my thesis examiners and finding many typos in the first version of the document (any remaining mistakes should certainly be attributed to me!).

I would like to acknowledge my research group, the Centre for Quantum Information and Foundations, which has hosted many interesting seminars and discussions. In particular, I am thankful for Sathyawageeswar Subramanian and Mithuna Yoganathan who started their PhDs at the same time as me, and have been great company on this journey.

I would also like to thank the Cantab Capital Institute for the Mathematics of Information (CCIMI) for the funding to undertake this PhD, and in particular Carola Schönlieb for her leadership and Rachel Furner and Josh Stevens for organizing everything. I’d also like to thank my cohort of students at the CCIMI: Edward Ayers, Andrew Celsus, Sam Power, Ferdia Sherry, and Sven Wang; it has been nice to share reading groups, offices, and houses with you. I feel like I have found a nice community, with others in Cambridge such as Lars Mennen, Federico Pasqualotto, and Sam Thomas, and with those further away, such as Peter Tsimiklis, Marshall Sumwalt, and Michail and Sarah Yasonik.

I also greatly appreciate the Julia programming language community, and in particular Alex Arslan, Benoît Legat, Harsha Nagarajan, and David P. Sanders. I am thankful to my masters supervisor Vojkan Jakšić who suggested to me that I apply to work with Nilanjana and Yan, and taught me many things. I also thank the administrative and facilities staff at the Centre for Mathematical Sciences and at Churchill College.

Lastly, I very much appreciate the support of my family, immediate and extended, and my partner, Hang Lu Su.



# Abstract

This thesis combines two parallel research directions: an exploration into the continuity properties of certain entropic quantities, and an investigation into a simple class of physical systems whose time evolution is given by the repeated application of a quantum channel.

In the first part of the thesis, we present a general technique for establishing local and uniform continuity bounds for Schur concave functions; that is, for real-valued functions which are decreasing in the majorization pre-order. Continuity bounds provide a quantitative measure of robustness, addressing the following question: If there is some uncertainty or error in the input, how much uncertainty is there in the output? Our technique uses a particular relationship between majorization and the trace distance between quantum states (or total variation distance, in the case of probability distributions). Namely, the majorization pre-order attains a maximum and a minimum over  $\varepsilon$ -balls in this distance. By tracing the path of the majorization-minimizer as a function of the distance  $\varepsilon$ , we obtain the path of “majorization flow”. An analysis of the derivatives of Schur concave functions along this path immediately yields tight continuity bounds for such functions.

In this way, we find a new proof of the Audenaert-Fannes continuity bound for the von Neumann entropy, and the necessary and sufficient conditions for its saturation, in a universal framework which extends to the other functions, including the Rényi and Tsallis entropies. In particular, we prove a novel uniform continuity bound for the  $\alpha$ -Rényi entropy with  $\alpha > 1$  with much improved dependence on the dimension of the underlying system and the parameter  $\alpha$  compared to previously known bounds. We show that this framework can also be used to provide continuity bounds for other Schur concave functions, such as the number of connected components of a certain random graph model as a function of the underlying probability distribution, and the number of distinct realizations of a random variable in some fixed number of independent trials as a function of the underlying probability mass function. The former has been used in modeling the spread of epidemics, while the latter has been studied in the context of estimating measures of biodiversity from observations; in these contexts, our continuity bounds provide quantitative estimates of robustness to noise or data collection errors.

In the second part, we consider repeated interaction systems, in which a system of interest interacts with a sequence of probes, i.e. environmental systems, one at a time. The state of the system after each interaction is related to the state of the system before the interaction by the so-called reduced dynamics, which is described by the action of a quantum channel. When each probe and the way it interacts with the system is identical, the reduced dynamics at each step is identical. In this scenario, under the additional assumption that the reduced dynamics satisfies a faithfulness property, we characterize which repeated interaction systems break any initially-present entanglement between the system and an untouched reference, after finitely many steps. In this case, the reduced dynamics is said to be *eventually entanglement-breaking*. This investigation helps improve our understanding of which kinds of noisy time evolution destroy entanglement.

When the probes and their interactions with the system are slowly-varying (i.e. adiabatic), we analyze the saturation of Landauer's bound, an inequality between the entropy change of the system and the energy change of the probes, in the limit in which the number of steps tends to infinity and both the difference between consecutive probes and the difference between their interactions vanishes. This analysis proceeds at a fine-grained level by means of a two-time measurement protocol, in which the energy of the probes is measured before and after each interaction. The quantities of interest are then studied as random variables on the space of outcomes of the energy measurements of the probes, providing a deeper insight into the interrelations between energy and entropy in this setting.

# Summary of contributions

Part I of this thesis combines the results of [HD18; HD19; AHD19; HKDW20], along with some as-of-yet unpublished work, namely Sections 3.4.3, 3.7, 5.1.1, 5.7 and 6.2. Chapter 3 and Chapter 7 are largely the results of [HD18], with is joint work with Nilanjana Datta, Chapter 4 and many of the continuity bounds of Chapter 5 are the results of [HD19], which is also joint work with Nilanjana Datta, Section 5.5 is the result of [AHD19], which is joint work with Koenraad Audenaert and Nilanjana Datta, while Chapter 6 contains many of the results of [HKDW20], which is joint work with Vishal Katariya, Nilanjana Datta, and Mark Wilde. Part II of this thesis combines some of the results of [HRS20], which is joint work with Cambyse Rouzé and Daniel Stilck França, in Chapter 11 and some of the results of [HJPR18], which is joint work with Alain Joye, Yan Pautrat, and Renaud Raquépas, in Chapter 12.

Thus, almost all of the results of this thesis were developed in collaboration with others. The proofs included in Chapters 1–11, however, are my own work, written with advice and suggestions from my supervisors Dr. Datta and Dr. Pautrat, and my other coauthors, with the exception of the proof of Theorem 5.5.2, which was largely constructed by Dr. Audenaert<sup>1</sup>. Chapter 12 provides a summary of the paper [HJPR18]; the few proofs included in that chapter were written collaboratively with my coauthors.

The papers [CHDH18b; CHDH18a], which are joint work with Hao-Chung Cheng, Nilanjana Datta, and Min-Hsiu Hsieh, were written during my PhD, but were omitted from this thesis in order to focus on the two themes of entropic continuity bounds and eventually entanglement-breaking channels.

[AHD19] Koenraad Audenaert, Eric P. Hanson, and Nilanjana Datta. “A Continuity Bound for the Expected Number of Connected Components of a Random Graph: A Model for Epidemics” (Dec. 11, 2019). arXiv: [1912.05599](https://arxiv.org/abs/1912.05599) [math].

---

<sup>1</sup>My contributions to that proof consist of the numerical bounds, reducing the task to bounding (5.46), and some small fixes.

- [CHDH18a] Hao-Chung Cheng, Eric P. Hanson, Nilanjana Datta, and Min-Hsiu Hsieh. “Duality between Source Coding with Quantum Side Information and C-q Channel Coding” (Sept. 28, 2018). arXiv: [1809.11143 \[quant-ph\]](#).
- [CHDH18b] Hao-Chung Cheng, Eric P. Hanson, Nilanjana Datta, and Min-Hsiu Hsieh. “Non-Asymptotic Classical Data Compression with Quantum Side Information” (Mar. 20, 2018). arXiv: [1803.07505 \[quant-ph\]](#).
- [HD18] Eric P. Hanson and Nilanjana Datta. “Maximum and Minimum Entropy States Yielding Local Continuity Bounds”. *Journal of Mathematical Physics* 59.4 (Apr. 1, 2018), p. 042204. DOI: [10.1063/1.5000120](#).
- [HD19] Eric P. Hanson and Nilanjana Datta. “Universal Proofs of Entropic Continuity Bounds via Majorization Flow” (Dec. 11, 2019). arXiv: [1909.06981 \[quant-ph\]](#).
- [HJPR18] Eric P. Hanson, Alain Joye, Yan Pautrat, and Renaud Raquépas. “Landauer’s Principle for Trajectories of Repeated Interaction Systems”. *Annales Henri Poincaré* 19.7 (July 1, 2018), pp. 1939–1991. DOI: [10.1007/s00023-018-0679-1](#).
- [HKDW20] Eric P. Hanson, Vishal Katariya, Nilanjana Datta, and Mark M. Wilde. “Guesswork with Quantum Side Information” (Feb. 3, 2020). arXiv: [2001.03598 \[quant-ph\]](#).
- [HRS20] Eric P. Hanson, Cambyse Rouzé, and Daniel Stilck França. “Eventually Entanglement Breaking Markovian Dynamics: Structure and Characteristic Times”. *Annales Henri Poincaré* (Mar. 11, 2020). DOI: [10.1007/s00023-020-00906-4](#).

# Table of contents

<b>I</b>	<b>Entropic continuity bounds</b>	<b>1</b>
<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Notation and definitions</b>	<b>7</b>
<b>3</b>	<b>Relationship between majorization and the 1-norm ball</b>	<b>11</b>
3.1	Quantum-to-classical reduction . . . . .	13
3.2	A first look at majorization over the 1-norm ball . . . . .	15
3.3	Constructing the majorization-minimizer (3.7) . . . . .	19
3.4	Three ways to find the majorization-minimizer . . . . .	21
3.4.1	From optimality conditions . . . . .	21
3.4.2	Waterfilling argument . . . . .	23
3.4.3	From majorization-infimum . . . . .	24
3.5	Minimality of $r_\varepsilon^*$ in the majorization order (3.7) . . . . .	31
3.5.1	Uniqueness of construction for $r_\varepsilon^*$ . . . . .	34
3.6	Constructing the maximal state in the majorization order (3.7) . . . . .	38
3.7	Other $p$ -norms . . . . .	41
<b>4</b>	<b>Majorization flow and continuity bounds</b>	<b>49</b>
4.1	Local bounds . . . . .	49
4.2	The majorization-minimizer map . . . . .	50
4.3	Majorization flow . . . . .	55
4.3.1	The majorization-minimizer for small $\varepsilon$ . . . . .	55
4.3.2	The path of majorization flow . . . . .	59
4.4	Uniform continuity bounds from majorization flow . . . . .	63
<b>5</b>	<b>Applications</b>	<b>67</b>
5.1	$(h, \phi)$ -entropies . . . . .	70
5.1.1	An aside on the necessity of the $(h, \phi)$ form . . . . .	76

5.2	Rényi and Tsallis entropies . . . . .	78
5.2.1	Rényi entropy of parameter $\alpha = 1$ has optimal dimensional scaling in its continuity bound . . . . .	85
5.2.2	Discussion of previous continuity bounds for $S_\alpha$ with $\alpha > 1$ . . . .	88
5.2.3	Connection to thermodynamics . . . . .	89
5.3	Other entropies . . . . .	91
5.4	Number of distinct realizations from $N$ i.i.d. random variables . . . . .	94
5.5	Number of connected components of a random graph . . . . .	96
5.5.1	Proof of the upper bound (5.42) . . . . .	100
5.5.2	Proof of the lower bound (5.43) . . . . .	104
5.5.3	Maximizing $x - x^2 M(x)$ . . . . .	106
5.6	Smoothed entropies and MaxEnt . . . . .	107
5.7	Application to LOCC transformations . . . . .	110
5.8	A continuity bound for $\rho \mapsto D(\rho  \sigma)$ . . . . .	113
<b>6</b>	<b>Guesswork</b>	<b>115</b>
6.1	The Shannon entropy in a binary guessing game . . . . .	115
6.2	The guesswork: A different guessing game . . . . .	118
6.3	The guesswork with quantum side information . . . . .	121
6.3.1	Statement of the problem . . . . .	124
6.3.2	Guessing strategies . . . . .	125
6.3.3	A semidefinite optimization representation . . . . .	133
6.3.4	Concavity and continuity of the guesswork with quantum side information . . . . .	136
6.3.5	Computing the guesswork . . . . .	139
6.3.6	Entropic bounds . . . . .	149
<b>7</b>	<b>General concave functions maximized over the <math>\varepsilon</math>-ball</b>	<b>157</b>
7.1	Tools from convex optimization . . . . .	157
7.2	Gâteaux-differentiable functions . . . . .	159
7.3	Optimality conditions for maximizing concave functions over the trace-ball	160
7.4	Proofs of the remaining lemmas and corollaries . . . . .	166
<b>8</b>	<b>Open questions from Part I</b>	<b>171</b>

<b>II</b>	<b>Eventually Entanglement-Breaking Channels</b>	<b>173</b>
<b>9</b>	<b>Introduction</b>	<b>175</b>
<b>10</b>	<b>Linear Algebra and entanglement theory</b>	<b>181</b>
10.1	Vectors, operators, and superoperators . . . . .	182
10.1.1	Vector spaces . . . . .	183
10.1.2	Operators on a vector space . . . . .	184
10.1.3	Operators on a Hilbert space . . . . .	187
10.1.4	Superoperators . . . . .	188
10.2	Separability and entanglement . . . . .	191
10.2.1	Separability at the level of operators . . . . .	192
10.2.2	Choi-Jamolkowski (CJ) isomorphism: translate between operators and superoperators . . . . .	193
10.2.3	EB maps: separability at the level of superoperators . . . . .	195
10.3	Peripheral spectral properties of superoperators . . . . .	197
<b>11</b>	<b>Characterizing EEB channels</b>	<b>209</b>
11.1	Introduction through (counter)examples . . . . .	209
11.2	Entanglement classes of quantum channels . . . . .	212
11.3	Which quantum channels are eventually entanglement-breaking? . . . . .	217
11.3.1	Irreducible evolutions . . . . .	218
11.3.2	Beyond irreducibility . . . . .	223
11.3.3	Reduction to the trace-preserving case . . . . .	225
<b>12</b>	<b>Repeated interaction systems</b>	<b>229</b>
12.1	Repeated interaction systems . . . . .	230
12.2	Two-time measurement protocols . . . . .	231
12.3	Landauer's Principle and the balance equation . . . . .	234
12.4	The balance equation at the level of trajectories . . . . .	236
12.5	Adiabatic limit of repeated interaction systems . . . . .	241
12.6	Moment generating functions and deformed CP maps . . . . .	243
12.7	An adiabatic theorem . . . . .	246
12.8	The statistics of the entropy production at the level of trajectories . . . . .	248
12.9	Examples . . . . .	253
<b>13</b>	<b>Open questions from Part II</b>	<b>259</b>

References	261
Appendix A Mathematica code	275



# Part I

## Entropic continuity bounds



# Chapter 1

## Introduction

Entropies play a fundamental role in classical and quantum information theory as characterizations of the optimal rates of information theoretic tasks, and as measures of uncertainty. The mathematical properties of entropic functions therefore have important physical implications. The von Neumann entropy  $S$ , given by

$$S(\rho) := -\operatorname{tr}[\rho \log \rho]$$

is strictly concave and continuous as a function of  $d$ -dimensional quantum states, and is bounded by  $\log d$ , where the logarithm is taken to be base 2. As the von Neumann entropy characterizes the optimal rate of data compression for a memoryless quantum information source [Sch96], continuity of the von Neumann entropy, for example, implies that the quantum data compression limit is continuous in the source state. Likewise, for probability distributions  $p \in \mathcal{P}_d$ , where  $\mathcal{P}_d$  is the set of probability distributions over  $\{1, \dots, d\}$ , the Shannon entropy,  $S(p) := -\sum_{i=1}^d p_i \log p_i$  characterizes the data compression limit of a classical i.i.d. source  $X \sim p$ . Note that the von Neumann entropy of a quantum state is given by the Shannon entropy of its eigenvalues. The Shannon entropy is commonly denoted by  $H(p)$ , but in this thesis we will reserve  $H$  for a generic single-partite entropic quantity.

The von Neumann entropy satisfies the following tight continuity bound: Given  $\varepsilon > 0$  and  $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ , where  $\mathcal{D}(\mathcal{H})$  is the set of density matrices over a Hilbert space  $\mathcal{H}$  with dimension  $d$ , with trace distance  $T(\rho, \sigma) := \frac{1}{2} \operatorname{tr} |\rho - \sigma| \leq \varepsilon$ , the following inequality holds:

$$|S(\rho) - S(\sigma)| \leq \begin{cases} \varepsilon \log(d-1) + h(\varepsilon) & \text{if } \varepsilon < 1 - \frac{1}{d} \\ \log d & \text{if } \varepsilon \geq 1 - \frac{1}{d} \end{cases} \quad (1.1)$$

where  $h(\varepsilon) := -\varepsilon \log \varepsilon - (1 - \varepsilon) \log(1 - \varepsilon)$  denotes the binary entropy. The same bound holds in the classical case for  $p, q \in \mathcal{P}_d$ , with total variation distance  $\text{TV}(p, q) := \frac{1}{2} \sum_{i=1}^d |p_i - q_i| \leq \varepsilon$ .

This inequality is known in the quantum information theory literature as the Audenaert-Fannes bound, which is a strengthened version of the Fannes bound [Fan73] and was established in [Aud07] by a direct optimization argument. The bound was also proven via a coupling argument in [Pet08, Theorem 3.8], with credit to Csiszár. In the classical case, it was also proven by [Zha07a, Cor. 1] by a coupling argument, and by Ho and Yeung [HY10, Theorem 6] via an optimization over local continuity bounds, similar in spirit to the techniques used later in this thesis (although with an analysis specific to the Shannon entropy, as opposed to more general families of entropies). The coupling argument was revisited in [Sas13, Theorem 3] and [Win16, Lemma 1]. In each case (except for [Pet08, Theorem 3.8]), it was shown that equality occurs if one state is pure, and the other state has the spectrum  $\{1 - \varepsilon, \frac{\varepsilon}{d-1}, \dots, \frac{\varepsilon}{d-1}\}$ . These conditions were shown to be necessary in [HD17] by an analysis of the coupling argument.

The results established in Part I of this thesis can largely be seen as investigations into, and generalizations of, the inequality (1.1) by interpreting it as a consequence of a relationship between trace distance and majorization.

Majorization [HLP29] (see also [Sch23] and [Hor54]) is an ordering of vectors (technically, a *preorder*) which provides a means to describe one vector as being more disordered, less disordered, or incomparable with another vector. Majorization has natural connections to entropies, which quantify disorder or randomness of probability distributions or quantum states, and thus provide a measure of information encoded in a random variable or a quantum state. In fact, almost every entropy considered in the information-theoretic literature is Schur concave, meaning that if one vector is more disordered than another, according to the majorization order, then it has higher entropy than the other. Perhaps as a consequence, majorization has proven to be a very useful tool in classical and quantum information theory.

Majorization can be extended from vectors to quantum states by simply defining that a quantum state  $\rho$  majorizes another quantum state  $\sigma$  if the vector of eigenvalues of  $\rho$  (counted with multiplicity) majorizes the vector of eigenvalues of  $\sigma$ . In quantum information theory, majorization plays a key role in the theory of bipartite pure state entanglement due to Nielsen's theorem [Nie99]. This result states that given two pure states  $\psi_{AB}$  and  $\phi_{AB}$  of a bipartite system  $AB$ , the state  $\psi_{AB}$  can be transformed into  $\phi_{AB}$  via local operations and classical communication if and only if the reduced state of  $\psi_{AB}$  for a subsystem ( $A$  or  $B$ ) is majorized by the corresponding reduced state of  $\phi_{AB}$ .

In Chapter 3 it is shown that the majorization order on probability vectors behaves well with respect to the total variation distance, in the sense that given any probability vector  $p$  and  $\varepsilon > 0$ , there exists a *minimal* and *maximal probability vector* in majorization order in the  $\varepsilon$ -ball around  $p$  with respect to the total variation distance. The existence of such majorization extrema has recently appeared in several forms. Both these probability vectors were independently found in [HOS18], and in fact the minimal vector already appears in [HY10]. The work [Kog13] developed related ideas in the context of balls in total-variation distance of subnormalized distributions centered at a normalized distribution, and the works [vdMee16] and [vdMNW17] consider a generalization of majorization called thermo-majorization, and find thermo-majorization extrema over balls in trace distance.

In the quantum setting, the existence of these states provides a natural connection between majorization and trace distance (and between majorization and the total variation distance in the classical setting). In Chapter 4, we develop a notion of *majorization flow*, which traces out the path of the minimizer in majorization order over the  $\varepsilon$ -ball as  $\varepsilon$  is changed infinitesimally.

Majorization flow turns out to be a very useful tool for obtaining remarkably simple and universal proofs of continuity bounds for numerous well-known families of entropies, including Rényi entropies [Rén61], Tsallis entropies [Tsa88], the so-called unified entropies [RT91], entropies induced by  $f$ -divergences [Pet85], and the concurrence [Woo01]. In particular, it allows us to establish Lipschitz continuity bounds for the  $\alpha$ -Rényi entropy with  $\alpha > 1$ . This resolves an open problem, left open since Audenaert's 2007 paper [Aud07], which was also presented at the Open Problem Session of the workshop *2017 Beyond I.I.D. in Information Theory* (held in Singapore), and provides a substantial improvement over previously known bounds.

**Overview of Part I** In the following, let  $H$  be a Schur concave function (such as the von Neumann entropy). Let us describe each chapter with an eye to how it provides insight into (1.1), or helps establish a generalization thereof.

- Chapter 2 introduces basic notation and definitions.
- Chapter 3 establishes the existence of a majorization-minimizer and -maximizer over the  $\varepsilon$ -ball in trace distance (or total variation distance, in the classical case). Hence the Schur concavity of  $H$  completely dictates its local continuity properties with respect to the trace distance, as manifested in Proposition 4.1.1.

- Chapter 4 considers the path of the majorization-minimizer over the  $\varepsilon$ -ball as  $\varepsilon$  varies, and shows that properties of the derivative  $\Gamma_H$  of  $H$  along the path provides insight into the global continuity properties of  $H$ . More specifically, a class of entropic quantities, which includes  $S$ , obtain a derivative  $\Gamma_H$  which itself is Schur convex. This implies that the value of “how much  $H$  can locally increase” itself respects the majorization order, and is maximized at pure states (or in the classical case,  $\delta$ -distributions such as  $p = (1, 0, \dots, 0)$ ). Since a state with spectrum  $(1 - \varepsilon, \frac{\varepsilon}{d-1}, \dots, \frac{\varepsilon}{d-1})$  (with multiplicity) obtains the maximum entropy over a trace-ball of radius  $\varepsilon$  centered at a pure state—unless the completely mixed state lies in this ball, in which case it obtains the maximum entropy over the ball—(1.1) follows.
- Chapter 5 applies this technique, and the related technique of bounding  $\Gamma_H$  to obtain a Lipschitz continuity bound, to obtain uniform continuity bounds for a range of other quantities: other entropies, including the Rényi and Tsallis entropies, the number of distinct observations from  $N$  independent trials, and the expected number of connected components of a particular model of random graph.
- Chapter 6 investigates the continuity of a quantity related to the Shannon entropy called the *guesswork*, with and without quantum side information. The guesswork without side information is not differentiable on all of  $\mathcal{P}_d$  but nevertheless a tight uniform continuity bound can be obtained with a slight generalization to the techniques previously addressed (although the quantity is simple enough that other techniques would likely suffice). The guesswork with side information, similarly to the conditional entropy, is not Schur concave, and the techniques described so far do not provide a uniform continuity bound. However, by analyzing a semidefinite optimization problem formulation of the quantity, a Lipschitz continuity bound can be obtained which is tight up to at most a multiplicative factor of 2 and an additive factor of 1.
- Chapter 7 calculates first-order optimality conditions for a quantum state to maximize a differentiable concave function over the  $\varepsilon$ -ball in trace distance around a given state. These conditions dictate the form of the majorization-minimizer discussed in Chapter 3 and provide necessary and sufficient conditions for a given state to maximize the quantum conditional entropy over the trace ball.

# Chapter 2

## Notation and definitions

We define the set of probability vectors of length  $d$  as  $\mathcal{P}_d \subseteq \mathbb{R}^d$ ,

$$\mathcal{P}_d := \left\{ p = (p_1, \dots, p_d) \in \mathbb{R}^d : p_i \geq 0 \text{ for } i = 1, \dots, d, \sum_{i=1}^d p_i = 1 \right\}.$$

We will often suppress the dependence on the dimension,  $\mathcal{P} \equiv \mathcal{P}_d$ .

The set of quantum states  $\mathcal{D}(\mathcal{H})$  where  $\mathcal{H}$  is a  $d$ -dimensional Hilbert space,  $d < \infty$ , is given by

$$\mathcal{D}(\mathcal{H}) := \{ \rho \in \mathcal{B}(\mathcal{H}) : \rho \geq 0, \text{tr}(\rho) = 1 \},$$

where  $\mathcal{B}(\mathcal{H})$  is the set of (bounded<sup>1</sup>) linear operators on  $\mathcal{H}$ . The set  $\mathcal{B}_{\text{sa}}(\mathcal{H}) \subseteq \mathcal{B}(\mathcal{H})$  is the set of self-adjoint operators on  $\mathcal{H}$ . The extremal elements of  $\mathcal{P}$  are permutations of the probability vector  $p := (1, 0, \dots, 0)$ . The extremal elements of  $\mathcal{D}(\mathcal{H})$  are rank-1 projections, and are called *pure states*. The identity matrix is denoted  $\mathbf{1} \equiv \mathbf{1}_{\mathcal{H}} \in \mathcal{B}(\mathcal{H})$ , where the dependence on  $\mathcal{H}$  in the notation may be suppressed.

We denote the completely mixed state by  $\tau := \frac{1}{d} \in \mathcal{D}(\mathcal{H})$ , and the analogous uniform distribution by  $u := (\frac{1}{d}, \dots, \frac{1}{d}) \in \mathcal{P}$ . A pure state is a rank-1 density matrix; we denote the set of pure states by  $\mathcal{D}_{\text{pure}}(\mathcal{H})$ . For two quantum states  $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ , the *trace distance* between them is given by

$$T(\rho, \sigma) := \frac{1}{2} \|\rho - \sigma\|_1$$

---

<sup>1</sup>In this thesis, we will only consider finite-dimensional Hilbert spaces, and hence all linear operators are bounded. We will thus slightly abuse notation and use  $\mathcal{B}(\mathcal{H})$  to denote the set of linear operators, the set of bounded operators, and the set of trace-class operators on  $\mathcal{H}$ , since all three sets coincide.

where  $\|A\|_1 = \text{tr} |A| = \sqrt{A^\dagger A}$  for  $A \in \mathcal{B}(\mathcal{H})$ . The trace distance  $T(\rho, \sigma)$  has an operational interpretation in terms of the optimal success probability  $p$  in distinguishing between two quantum states  $\rho$  and  $\sigma$  by a 2-outcome POVM, namely

$$p = \frac{1}{2}(1 + T(\rho, \sigma)).$$

Hence, the trace distance can be seen as a measure of distinguishability between  $\rho$  and  $\sigma$ . Analogously, the total variation distance between  $p, q \in \mathcal{P}$  is defined as

$$\text{TV}(p, q) := \frac{1}{2}\|p - q\|_1$$

and is analogously endowed with an interpretation in terms of distinguishability.

A function  $F : \mathcal{D}(\mathcal{H}) \rightarrow \mathbb{R}$  is *k-Lipschitz* (with respect to the trace distance) if for all  $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ ,

$$|F(\rho) - F(\sigma)| \leq k T(\rho, \sigma),$$

and likewise  $F : \mathcal{P} \rightarrow \mathbb{R}$  is *k-Lipschitz* if for all  $p, q \in \mathcal{P}$ ,  $|F(p) - F(q)| \leq k \text{TV}(p, q)$ . The smallest  $k > 0$  such that  $F$  is *k-Lipschitz* is called the *optimal Lipschitz constant* for  $F$ . The function  $F$  is said to be *Lipschitz continuous* if it is *k-Lipschitz* for some  $k > 0$ .

For  $\varepsilon > 0$ , we define the  $\varepsilon$ -ball (in trace distance) around  $\sigma \in \mathcal{D}(\mathcal{H})$  as the set

$$B_\varepsilon(\sigma) := \{\omega \in \mathcal{D}(\mathcal{H}) : T(\omega, \sigma) \leq \varepsilon\}, \quad (2.1)$$

and likewise the  $\varepsilon$ -ball (in total variation distance) around a probability vector  $p \in \mathcal{P}$  as the set

$$B_\varepsilon(p) := \{q \in \mathcal{P} : \text{TV}(p, q) \leq \varepsilon\}. \quad (2.2)$$

For any  $A \in \mathcal{B}_{\text{sa}}(\mathcal{H})$ , let  $\lambda_+(A)$  and  $\lambda_-(A)$  denote the maximum and minimum eigenvalue of  $A$ , respectively, and  $k_+(A)$  and  $k_-(A)$  denote their multiplicities. Let  $\lambda_j^\downarrow(A)$  denote the  $j$ th largest eigenvalue, counting multiplicity; that is, the  $j^{\text{th}}$  element of the ordering

$$\lambda_1^\downarrow(A) \geq \lambda_2^\downarrow(A) \geq \cdots \geq \lambda_d^\downarrow(A).$$

We set  $\vec{\lambda}^\downarrow(A) := (\lambda_i^\downarrow(A))_{i=1}^d \in \mathbb{R}^d$  and denote the spectrum of  $A \in \mathcal{B}_{\text{sa}}(\mathcal{H})$  (i.e. its set of eigenvalues) by  $\text{spec } A \subseteq \mathbb{R}$ .

The set of probability vectors with strictly positive entries is denoted  $\mathcal{P}_+$ . For a vector  $r \in \mathbb{R}^d$ ,  $r_+$  denotes its largest entry, and  $r_-$  denotes its smallest entry, and  $k_+$  the multiplicity of the largest entry  $r_+$  in  $r$ , and likewise  $k_-$  the multiplicity of the smallest entry. For  $r \in \mathbb{R}^d$ , we denote  $r_i^\downarrow$  as the  $i$ th largest entry of  $r$  (counting multiplicity), and



let  $r^\downarrow := (r_i^\downarrow)_{i=1}^d$ . Analogously,  $r_i^\uparrow$  denotes the  $i$ th smallest entry of  $r$ , counting multiplicity, and  $r^\uparrow := (r_i^\uparrow)_{i=1}^d$ . We set

$$\mathcal{P}^\downarrow \equiv \mathcal{P}_d^\downarrow := \{q^\downarrow : q \in \mathcal{P}\} = \{q \in \mathcal{P} : q_1 \geq q_2 \geq \dots \geq q_d\}.$$

We use  $\log x$  to denote the base-2 logarithm of  $x$  and  $\ln x$  to denote the natural logarithm of  $x$ .

**Majorization of vectors** For  $x, y \in \mathbb{R}^d$ , we say  $x$  *majorizes*  $y$ , written  $x \succ y$ , if

$$\sum_{j=1}^k x_j^\downarrow \geq \sum_{j=1}^k y_j^\downarrow \quad \forall k = 1, \dots, d-1, \quad \text{and} \quad \sum_{j=1}^d x_j^\downarrow = \sum_{j=1}^d y_j^\downarrow. \quad (2.3)$$

We have that  $x \succ y$  if and only if

$$y = Mx \quad (2.4)$$

for some doubly-stochastic matrix  $M$  [MOA11, Section 2, B.2]. We say a function  $\varphi : \mathcal{P} \rightarrow \mathbb{R}$  is Schur convex on a set  $S \subseteq \mathcal{P}$  if for  $p, q \in S$ ,  $p \prec q \implies \varphi(p) \leq \varphi(q)$ . If  $S = \mathcal{P}$ , we simply say  $\varphi$  is Schur convex. We say  $\varphi$  is *Schur concave* on  $S$  if  $-\varphi$  is Schur convex on  $S$ , and likewise,  $\varphi$  is Schur concave if  $-\varphi$  is Schur convex. One useful characterization of Schur convex functions is if  $\varphi : \mathcal{P} \rightarrow \mathbb{R}$  is symmetric and continuous, and differentiable on  $\mathcal{P}_+$ , then it is Schur convex if and only if

$$(p_i - p_j) [\partial_{p_i} \varphi(p) - \partial_{p_j} \varphi(p)] \geq 0 \quad \forall i, j \quad (2.5)$$

for each  $p \in \mathcal{P}$  [MOA11, Section 3, A.4.a].

**Majorization of quantum states** Given two quantum states  $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ , we say  $\sigma$  majorizes  $\rho$ , written  $\rho \prec \sigma$  if  $\vec{\lambda}^\downarrow(\rho) \prec \vec{\lambda}^\downarrow(\sigma)$ . We say that  $\varphi : \mathcal{D}(\mathcal{H}) \rightarrow \mathbb{R}$  is *Schur convex* if  $\varphi(\rho) \leq \varphi(\sigma)$  for any  $\rho, \sigma \in \mathcal{D}(\mathcal{H})$  with  $\rho \prec \sigma$ . If  $\varphi(\rho) < \varphi(\sigma)$  for any  $\rho, \sigma \in \mathcal{D}(\mathcal{H})$  such that  $\rho \prec \sigma$ , and  $\rho$  is not unitarily equivalent to  $\sigma$ , then  $\varphi$  is *strictly Schur convex*. We say  $\varphi$  is Schur concave (resp. strictly Schur concave) if  $(-\varphi)$  is Schur convex (resp. strictly Schur convex).



## Chapter 3

# Relationship between majorization and the 1-norm ball

This chapter is concerned with an interesting relationship between the majorization pre-order and the 1-norm: namely, that balls in 1-norm admit a majorization-*maximum* and -*minimum*, despite not all elements of the ball being pairwise comparable. More precisely, let  $r \in \mathcal{P}$  be a probability vector, and consider the total-variation ball

$$B_\varepsilon(r) := \left\{ q \in \mathcal{P} : \text{TV}(r, q) := \frac{1}{2} \|r - q\|_1 \leq \varepsilon \right\}$$

for  $\varepsilon > 0$ . There exists  $r_\varepsilon^*, r_{*,\varepsilon} \in B_\varepsilon(r)$  such that

$$r_\varepsilon^* \prec q \prec r_{*,\varepsilon} \quad \forall q \in B_\varepsilon(r). \quad (3.1)$$

We call  $r_\varepsilon^*$  the *majorization-minimizer* over  $B_\varepsilon(r)$ , and  $r_{*,\varepsilon}$  the *majorization-maximizer* over  $B_\varepsilon(r)$ . We place the upper star on the majorization-minimizer as we derived the states in [HD18] in the context of maximizing Schur concave functions.

In the following, we will discuss some of the recent history of (3.1), and how that history connects with the focus of this chapter. The first steps towards establishing (3.1) begin in Section 3.1.

This relationship between majorization and total variation distance, or key parts thereof, was independently found on at least three occasions: by Ho and Yeung in 2010 [HY10], by my supervisor and I in 2018 [HD18], and by Horodecki, Oppenheim, and Sparaciari in 2018 [HOS18]<sup>1</sup>. A generalization of these states were also constructed

---

<sup>1</sup>Horodecki and Oppenheim had known about this relationship for several years already, however [OH17].

in the context of thermal majorization by van der Meer and Wehner [vdMee16], and subsequently used to study approximate state transitions in [vdMNW17].

As an immediate consequence of (3.1), if  $H$  is Schur concave, then  $r_\varepsilon^*$  maximizes  $H$  over  $B_\varepsilon(r)$ :

$$r_\varepsilon^* \in \operatorname{argmax}_{q \in B_\varepsilon(r)} H(q)$$

This is used in Section 5.6 for calculating smoothed entropies, and provides a method for obtaining  $r_\varepsilon^*$  by optimization.

In this chapter, we present three ways to obtain  $r_\varepsilon^*$ . The first, based on [HD18] and presented in Section 3.4.1, uses necessary and sufficient conditions for a given state to be a maximizer of some concave function  $H$  (i.e. not necessarily unitarily invariant) over a trace distance ball in  $\mathcal{D}(\mathcal{H})$  (established in Chapter 7) by specializing to the case in which  $H$  is the von Neumann entropy. The general conditions reduce to ones suggestive of a particular form, which turns out to be the form of  $\operatorname{diag}(r_\varepsilon^*)$ . In fact, these “reduced conditions” can be obtained directly from a waterfilling argument applied to maximizing the Shannon entropy over a total-variation distance ball as shown in Section 3.4.2. The state can also be obtained by calculating the majorization-infimum over the ball and checking that it indeed resides inside the ball itself. This is discussed in Section 3.4.3.

To obtain the majorization-maximizer  $r_{*,\varepsilon}$ , which turns out to have a very simple construction, one simply uses that  $B_\varepsilon(r)$  is a polytope, and  $r_{*,\varepsilon}$  is necessarily one of its vertices (see Section 3.2 for more).

[HY10] constructed the majorization-minimizer  $r_\varepsilon^*$ , and [HOS18] constructed both the minimizer and maximizer, but neither discuss in detail how they were obtained; possibly, a waterfilling argument like the one sketched in Section 3.4.2 was used.

The relation (3.1) immediately provides *local* continuity bounds, in the sense that if  $H : \mathcal{P} \rightarrow \mathbb{R}$  is a Schur concave function, and  $r \in \mathcal{P}$  is given, then the variation in  $H$  near  $r$  can be bounded as

$$|H(q) - H(r)| \leq \max \{H(r) - H(r_{*,\varepsilon}), H(r_\varepsilon^*) - H(r)\} \quad \forall q \in B_\varepsilon(r). \quad (3.2)$$

These bounds, the relation (3.1), and the derivation of optimality conditions for maximizing a concave function over a trace-ball, were the focus of [HD18]. In that work, we also proved a semi-group property, namely for  $r \in \mathcal{P}$  and  $\varepsilon_1, \varepsilon_2 > 0$ ,

$$r_{\varepsilon_1 + \varepsilon_2}^* = (r_{\varepsilon_1}^*)_{\varepsilon_2}^*. \quad (3.3)$$

This property plays a central role in Chapter 4.

The work [HY10] instead focussed on establishing a *uniform* continuity bound (in contrast to a local one), for the Shannon entropy. In fact, likely unbeknownst to the authors, the optimal uniform continuity bound had already been established in 2007 by Audenaert [Aud07] and independently by [Zha07b], as discussed in Chapter 1. [HY10] derive this bound in a different manner, with a proof prominently featuring majorization, but also using heavily the specific expression for the Shannon entropy (as opposed to a more general Schur concave function).

[HOS18] established the existence of the majorization-minimizer and -maximizer states and used them to define a “majorization distance” between probability distributions, which is used in Section 5.7. They also proved that if  $p \prec q$ , then  $p_\varepsilon^* \prec q_\varepsilon^*$  for all  $\varepsilon > 0$ , which will prove useful in establishing uniform continuity bounds (and appears in Theorem 4.2.1).

In [HD17], my supervisor and I showed that to establish a uniform continuity bound for a certain class of Schur concave functions, one only needs to establish the bound for  $\varepsilon$  small enough, and then use the semigroup property finitely many times to establish the bound for all  $\varepsilon > 0$ . While this provided a new technique for establishing uniform continuity bounds, most of the applicable entropic quantities already had known tight uniform continuity bounds. The biggest limitation of the technique is that it only applied to concave functions. However, one of the most widely used entropies without a known tight uniform continuity bound was the  $\alpha$ -Rényi entropy

$$S_\alpha(r) = \frac{1}{1-\alpha} \log \sum_i r_i^\alpha$$

for  $\alpha > 1$ , which is neither concave nor convex (but is Schur concave). In trying to establish a uniform continuity bound for  $S_\alpha$ , eventually it became apparent that an infinitesimal approach provided a much simpler way to prove the result. The semigroup property naturally allows statements about  $r_\varepsilon^*$  to be derived in terms of  $\partial_\varepsilon r_\varepsilon^*|_{\varepsilon=0}$ , leading to the notion of a *path of majorization flow*, described in Chapter 4 (and the preprint [HD19]).

### 3.1 Quantum-to-classical reduction

In this chapter, we are interested in the relationship between the majorization of quantum states and the trace distance, with an interest in developing continuity bounds for Schur concave functions on the set of density matrices. This task reduces to a “classical” one: understanding the relationship between the majorization of probability vectors and

the total variation distance, and applying that to continuity bounds for Schur concave functions on the probability simplex.

First, the reduction for majorization is trivial:  $\rho \prec \sigma$  if and only if  $\vec{\lambda}^\downarrow(\rho) \prec \vec{\lambda}^\downarrow(\sigma)$ , by definition. Next, let us consider the trace distance. Let  $A \in \mathcal{B}_{\text{sa}}(\mathcal{H})$ . Then let  $\text{Eig}^\downarrow(A) := \text{diag}(\vec{\lambda}^\downarrow(A))$  be the diagonal matrix whose diagonal entries are the sorted eigenvalues of  $A$ , counted with multiplicity. By [Bha97, p. IV.62], we have for  $A, B \in \mathcal{B}_{\text{sa}}(\mathcal{H})$

$$\|\|\text{Eig}^\downarrow(A) - \text{Eig}^\downarrow(B)\|\| \leq \|A - B\| \quad (3.4)$$

for any unitarily invariant norm  $\|\|\cdot\|\|$ . In the case of the trace distance,

$$T(\text{Eig}^\downarrow(A), \text{Eig}^\downarrow(B)) \equiv \frac{1}{2} \|\text{Eig}^\downarrow(A) - \text{Eig}^\downarrow(B)\|_1 = \frac{1}{2} \|\vec{\lambda}^\downarrow(A) - \vec{\lambda}^\downarrow(B)\|_1 \equiv \text{TV}(\vec{\lambda}^\downarrow(A), \vec{\lambda}^\downarrow(B))$$

is the total variation distance between the sorted vectors of eigenvalues of  $A$  and  $B$ , and (3.4) yields

$$\text{TV}(\vec{\lambda}^\downarrow(A), \vec{\lambda}^\downarrow(B)) \leq T(A, B). \quad (3.5)$$

I first encountered this reduction in [Aud07].

Lastly, let  $H : \mathcal{D}(\mathcal{H}) \subseteq \mathcal{B}_{\text{sa}}(\mathcal{H}) \rightarrow \mathbb{R}$  be a unitarily invariant function. Then  $H(A) = H(\text{Eig}^\downarrow(A))$  is a function of the eigenvalues of  $A$  alone:  $H(A) = H_{\text{cl}}(\vec{\lambda}^\downarrow(A))$  for some function  $H_{\text{cl}} : \mathcal{P} \rightarrow \mathbb{R}$  (namely  $H_{\text{cl}}(r) := H(\text{diag}(r))$  for  $r \in \mathcal{P}$ , where  $\text{diag}(r)$  is the diagonal matrix with the entries of  $r$  upon its main diagonal).

We can see that these reductions interact favorably for establishing continuity bounds, in the following sense. If  $\frac{1}{2} \|\rho - \sigma\|_1 \leq \varepsilon$ , then  $\frac{1}{2} \|\vec{\lambda}^\downarrow(\rho) - \vec{\lambda}^\downarrow(\sigma)\|_1 \leq \varepsilon$  as well, and

$$|H(\rho) - H(\sigma)| = |H_{\text{cl}}(\vec{\lambda}^\downarrow(\rho)) - H_{\text{cl}}(\vec{\lambda}^\downarrow(\sigma))|.$$

Hence, it remains to bound this difference in terms of  $\varepsilon$ , using  $\frac{1}{2} \|\vec{\lambda}^\downarrow(\rho) - \vec{\lambda}^\downarrow(\sigma)\|_1 \leq \varepsilon$ . This is precisely the task of establishing a continuity bound on  $\mathcal{P}$  with respect to the total variation distance.

Moreover, this reduction preserves tightness, as shown by the following simple argument. Assume we have established a bound of the form

$$|H_{\text{cl}}(p) - H_{\text{cl}}(q)| \leq f(\varepsilon)$$

for all  $p, q \in \mathcal{P}$  such that  $\text{TV}(p, q) \leq \varepsilon$ , for some function  $f$ . Then by the above reduction, we know that for all  $\rho, \sigma \in \mathcal{D}(\mathcal{H})$  with  $T(\rho, \sigma) \leq \varepsilon$ ,

$$|H(\rho) - H(\sigma)| \leq f(\varepsilon).$$

If two vectors  $p^*, q^* \in \mathcal{P}$  with  $\text{TV}(p^*, q^*) \leq \varepsilon$  achieve  $|H_{\text{cl}}(p^*) - H_{\text{cl}}(q^*)| = f(\varepsilon)$ , then  $\rho^* := \text{diag}(p^*)$  and  $\sigma^* := \text{diag}(q^*)$  satisfy

$$|H(\rho^*) - H(\sigma^*)| = f(\varepsilon).$$

That is, a tight bound on the classical level yields a tight bound on the quantum level.

## 3.2 A first look at majorization over the 1-norm ball

Fix  $\varepsilon > 0$  and a state  $\rho \in \mathcal{D}(\mathcal{H})$ , with  $d := \dim \mathcal{H}$ . The main result of this chapter is that the  $\varepsilon$ -ball around  $\rho$ , the  $\varepsilon$ -ball in trace distance,  $B_\varepsilon(\rho)$ , admits a minimum and maximum in the majorization order. Note that since majorization is a partial order, a priori one does not know that there are states in  $B_\varepsilon(\rho)$  comparable to every other state in  $B_\varepsilon(\rho)$ .

**Theorem 3.2.1.** *Let  $\rho \in \mathcal{D}$  and  $\varepsilon > 0$ . Then there exist two states  $\rho_\varepsilon^*$  and  $\rho_{*,\varepsilon}$  in  $B_\varepsilon(\rho)$  (which are defined by Equation (3.16) in Section 3.3 and Equation (3.42) in Section 3.6 respectively) both of which commute with  $\rho$  and satisfy*

$$\rho_\varepsilon^* \prec \omega \prec \rho_{*,\varepsilon}, \quad \forall \omega \in B_\varepsilon(\rho). \quad (3.6)$$

Moreover,  $\rho_\varepsilon^*$  is the unique state in  $B_\varepsilon(\rho)$  satisfying the left-hand relation of (3.6), and  $\rho_{*,\varepsilon}$  is the unique state in  $B_\varepsilon(\rho)$  satisfying the right-hand relation of (3.6) up to unitary equivalence.

This result can be specialized to the case of probability vectors. Let  $p \in \mathcal{P}_d$  and  $\varepsilon > 0$ . Then there exist two probability vectors  $p_\varepsilon^*$  and  $p_{*,\varepsilon}$  in  $B_\varepsilon(p)$  (which are defined by Equation (3.13) in Section 3.3 and Equation (3.40) in Section 3.6 respectively) which satisfy

$$p_\varepsilon^* \prec q \prec p_{*,\varepsilon}, \quad \forall q \in B_\varepsilon(p). \quad (3.7)$$

Moreover,  $p_\varepsilon^*$  is the unique element of  $B_\varepsilon(p)$  satisfying the left-hand relation of (3.7), and  $p_{*,\varepsilon}$  is the unique element of  $B_\varepsilon(p)$  satisfying the right-hand relation of (3.7) up to permutations.

*Remark.* See Figure 3.1 for an example of  $\rho_\varepsilon^*$  and  $\rho_{*,\varepsilon}$  for a particular state  $\rho$ .

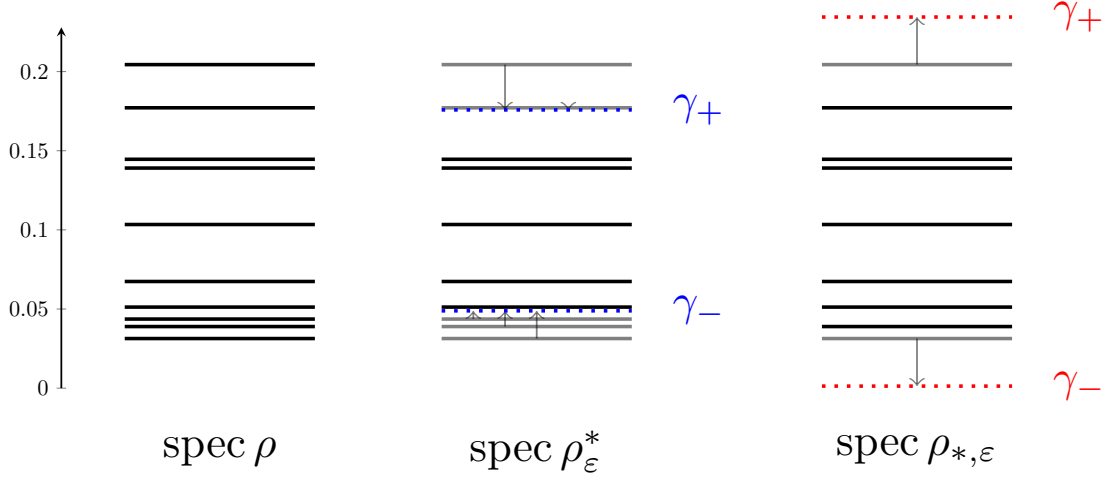


Fig. 3.1 An example of the majorization-minimizer and majorization-maximizer with  $d = 2$ . The spectrum of a quantum state  $\rho$  is on the left, the spectra of  $\rho_\varepsilon^*$  in the center and  $\rho_{*,\varepsilon}$  on the right, with  $\varepsilon = 0.03$ . The states are constructed in Section 3.3 and Section 3.6 respectively, and have the structure of raising or lowering some eigenvalues of  $\rho$  according to a prescribed formula.

In this section, we consider the  $\varepsilon$ -ball  $B_\varepsilon(\rho)$  around a state  $\rho \in \mathcal{D}(\mathcal{H})$ , and motivate the construction of maximal and minimal states in the majorization order given in (3.6). As discussed in Section 3.1, we prove Theorem 3.2.1 by reducing it to the classical case of discrete probability distributions on  $d$  symbols, and then constructing explicit states  $\rho_\varepsilon^*$  (in Section 3.3), and  $\rho_{*,\varepsilon}$  (in Section 3.6), whose eigenvalues are respectively given by the probability distributions which are minimal and maximal in majorization order.

As discussed in Section 3.1, it suffices to consider the simplex of probability vectors  $\mathcal{P}_d$ , equipped with the total variation distance, instead of the set of density operators  $\mathcal{D}(\mathcal{H})$  equipped with the trace distance. Note that  $\mathcal{P}_d$  is the polytope (i.e. the convex hull of finitely many points) generated by  $(1, 0, \dots, 0)$  and its permutations. Recall the total variation ball,

$$B_\varepsilon(p) = \{w = (w_i)_{i=1}^d \in \mathcal{P}_d : \frac{1}{2} \|w - p\|_1 \leq \varepsilon\}.$$

The set  $\{x \in \mathbb{R}^d : \|x\|_1 \leq 1\}$  can be written

$$\{x \in \mathbb{R}^d : \|x\|_1 \leq 1\} = \text{conv}\{e_1, -e_1, \dots, e_d, -e_d\},$$

where  $\text{conv}(A)$  denotes the convex hull of a set  $A$ , and  $e_1, \dots, e_d$  are the vectors of the standard basis (e.g.  $e_j = (0, \dots, 0, 1, 0, \dots, 0)$  with 1 in the  $j$ th entry), and is therefore a



polytope, called the  $d$ -dimensional cross-polytope (see e.g. [Mat02, p. 82]). As a translation and scaling of the  $d$ -dimensional cross-polytope, the set  $\{w \in \mathbb{R}^d : \frac{1}{2}\|w - p\|_1 \leq \epsilon\}$  is a polytope as well. As  $B_\epsilon(p)$  is the intersection of this set and  $\mathcal{P}_d$ , it too is a polytope. See Figure 3.2a for an illustration of  $B_\epsilon(p)$  in a particular example.

The existence of  $p_\epsilon^*$  and  $p_{*,\epsilon}$  in  $B_\epsilon(p)$  satisfying (3.7) is equivalent to  $p_\epsilon^*$  and  $p_{*,\epsilon}$  in  $B_\epsilon(p)$  satisfying

$$p_\epsilon^* \prec w \prec p_{*,\epsilon} \quad (3.8)$$

for all  $w \in B_\epsilon(p)$ . Using Birkhoff's Theorem (e.g. [NC09, Theorem 12.12]), the set of vectors majorized by a point  $w \in \mathcal{P}_d$  can be shown to be given by

$$M_w := \{p \in \mathcal{P}_d : w \succ p\} = \text{conv}\{\pi(w) : \pi \in S_d\}, \quad (3.9)$$

where  $S_d$  is the symmetric group on  $d$  letters (see [MOA11, p. 34]). Let us illustrate this with an example in  $d = 3$ . Let us choose  $p = (0.21, 0.24, 0.55)$  and  $\epsilon = 0.1$ . The simplex  $\mathcal{P}_d$  and ball  $B_\epsilon(p)$  are depicted in Figure 3.2a. A point  $w = (0.14, 0.28, 0.58) \in B_\epsilon(p)$  is shown in Figure 3.2b, and the set  $M_w$  in Figure 3.2c.

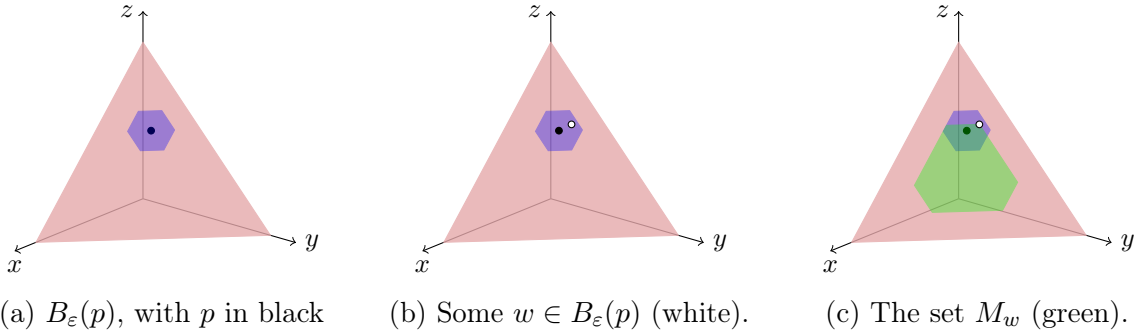


Fig. 3.2 In dimension  $d = 3$ , the simplex,  $\mathcal{P}_d$ , of probability vectors is the shaded triangle shown in (a), along with the ball  $B_\epsilon(p)$  which is the hexagon shown in blue, centered at  $p = (0.21, 0.24, 0.55)$  (depicted by a black dot) with  $\epsilon = 0.1$ . In (b), a point  $w = (0.14, 0.28, 0.58)$  is depicted in white, and in (c), the set  $M_w$  is shown in green.

The geometric characterization (3.9), depicted in Figure 3.2, requires that  $B_\epsilon(p) \subseteq M_{p_{*,\epsilon}}$ , and conversely,  $p_\epsilon^* \in M_p$  for each  $p \in B_\epsilon(p)$ . Figure 3.2c shows that for the point  $w$ ,  $B_\epsilon(p) \not\subseteq M_w$ , implying that  $w \neq p_{*,\epsilon}$ . Moreover, one can check that that e.g.  $w \notin M_p$ , and hence  $w \neq p_\epsilon^*$ .

Next we consider Schur concave functions on  $\mathcal{P}_d$ , in order to gain insight into the probability distributions  $p_{*,\epsilon}$  and  $p_\epsilon^*$  which arise in the majorization order (3.8). In particular, let us consider Shannon entropy  $S(w) := -\sum_{i=1}^d w_i \log w_i$  of a probability distribution  $w = (w_i)_{i=1}^d$ . It is known to be strictly Schur concave. Hence, if  $p_{*,\epsilon} \in B_\epsilon(p)$

satisfying (3.8) exists, it must satisfy

$$S(p_{*,\epsilon}) \leq S(w)$$

for any  $w \in B_\epsilon(p)$ . Thus,  $p_{*,\epsilon}$  must be a minimizer of  $S$ , which is a concave function, over  $B_\epsilon(p)$ , a convex set. Similarly,  $p_\epsilon^*$  must be a maximizer of  $S$  over  $B_\epsilon(p)$ . Properties of maximizers of concave functions over a convex sets are well-understood; in particular, any local maximizer is a global maximizer.

The task of minimizing a concave function over a convex set is a priori more difficult; in particular, local minima need not be global minima. There is, however, a *minimum principle* which asserts that the minimum occurs on the boundary of the set; this is formulated more precisely in e.g. [Roc96, Chapter 32]. Since  $B_\epsilon(p)$  is a polytope,  $S$  is minimized on one of the finitely many vertices of  $B_\epsilon(p)$ . This fact yields a simple solution to the problem of minimizing  $S$  over  $B_\epsilon(p)$ , as described below by example, and in generality in Section 3.6.

Let us return to the example of Figure 3.2. We see  $B_\epsilon(p)$  has six vertices; these are  $\{p + \pi((\epsilon, -\epsilon, 0)) : \pi \in S_d\}$ . The vertex which minimizes  $S$  is  $v := (0.21 - \epsilon, 0.24, 0.55 + \epsilon)$ ,



(a) A minimum  $v$  of  $S$  over  $B_\epsilon(p)$ , in white. (b) Maximum  $m$  of  $S$  over  $B_\epsilon(p)$ , in white.

Fig. 3.3 For the example of Figure 3.2, the (unique) maximum and minimum of the Shannon entropy  $S$  over  $B_\epsilon(p)$  are shown. Both  $v$  and  $m$  occur on the boundary of  $B_\epsilon(p)$ .

where the smallest entry is decreased and the largest entry is increased, as shown in Figure 3.3b. Moreover, one can check that  $w \prec v$  for any  $w \in B_\epsilon(p)$ . This leads us to the conjecture that the vertex corresponding to decreasing the smallest entry and increasing the largest entry will yield  $p_{*,\epsilon}$  satisfying (3.8) in general. We see in Section 3.6 that this is indeed true, although in some cases more than one entry needs to be decreased.

On the other hand, finding the probability distribution  $p_\epsilon^*$  in  $B_\epsilon(p)$  which satisfies (3.8) is more than a matter of checking the vertices of  $B_\epsilon(p)$ , as shown by Figure 3.3b: in this example,  $p_\epsilon^*$  is not a vertex of  $B_\epsilon(p)$ . Interestingly, useful insight into this probability

distribution can be obtained by using results from convex optimization theory. This is discussed in the following section.

### 3.3 Constructing the majorization-minimizer (3.7)

Let  $r \in \mathcal{P}_d$  and  $\varepsilon > 0$ . If  $u \in B_\varepsilon(r)$ , then  $r_\varepsilon^* = u$ , where  $u := (1/d, \dots, 1/d)$  denotes the uniform distribution, so instead consider the case in which  $u \notin B_\varepsilon(r)$ . We first define a quantity  $\gamma_+^{(m)} \equiv \gamma_+^{(m)}(r, \varepsilon)$ , for  $m \in \{1, \dots, d-1\}$ , as

$$\gamma_+^{(m)} := \frac{1}{m} \left( \sum_{i=1}^m r_i^\downarrow - \varepsilon \right).$$

Similarly, define  $\gamma_-^{(m)} \equiv \gamma_-^{(m)}(r, \varepsilon)$  by

$$\gamma_-^{(m)} := \frac{1}{m} \left( \sum_{i=1}^m r_i^\uparrow + \varepsilon \right) = \frac{1}{m} \left( \sum_{i=d-m+1}^d r_i^\downarrow + \varepsilon \right).$$

Then we define  $m_+ := m_+(r, \varepsilon)$  as the unique solution to the following inequalities:

$$r_{m+1}^\downarrow \leq \gamma_+^{(m)} < r_m^\downarrow, \quad m \in \{1, \dots, d-1\} \quad (3.10)$$

and we set  $m_+(u, \varepsilon) = 0$ . Similarly, for  $r \neq u$ , we define  $m_- := m_-(r, \varepsilon)$  as the unique solution to the inequalities:

$$r_m^\uparrow < \gamma_-^{(m)} \leq r_{m+1}^\uparrow, \quad m \in \{1, \dots, d-1\} \quad (3.11)$$

or equivalently, the inequalities

$$r_{d-m+1}^\downarrow < \gamma_-^{(m)} \leq r_{d-m}^\downarrow, \quad m \in \{1, \dots, d-1\}. \quad (3.12)$$

Lastly, we set  $\gamma_+ = \gamma_+(r, \varepsilon) := \gamma_+^{(m_+)}$  and  $\gamma_- = \gamma_-(r, \varepsilon) := \gamma_-^{(m_-)}$ . The proof that the above equations have unique solutions follows from Lemma 3.5.2, which is postponed to the end of this section. In the following, we use Dirac notation where for  $i = 1, \dots, d$ , the symbol  $|i\rangle$  denotes the vector  $e_i = (0, \dots, 0, 1, 0, \dots, 0)$ , with the 1 in the  $i$ th place. Then, given  $r = \sum_{i=1}^d r_i^\downarrow |\pi(i)\rangle$  for some permutation  $\pi \in S_d$  and where  $\{|i\rangle\}_{i=1}^d$  is the

standard basis of  $\mathbb{R}^d$ , we define

$$r_\varepsilon^* := \sum_{i=1}^{m_+} \gamma_+ |\pi(i)\rangle + \sum_{i=m_++1}^{d-m_-} r_i^\downarrow |\pi(i)\rangle + \sum_{i=d-m_-+1}^d \gamma_- |\pi(i)\rangle. \quad (3.13)$$

Note that for any permutation matrix  $P$  we have

$$(Pr)_\varepsilon^* = P(r_\varepsilon^*). \quad (3.14)$$

We check that  $r_\varepsilon^* \in B_\varepsilon(r)$  as follows. From its definition, its entries lie in the interval  $(0, 1]$  and  $r_\varepsilon^*$  has sum 1. Additionally, if  $\text{TV}(r, u) \leq \varepsilon$ , then  $r_\varepsilon^* = u \in B_\varepsilon(r)$ . Otherwise,

$$\|r_\varepsilon^* - r\|_1 = \sum_{i=1}^{m_+} |\gamma_- - r_i^\downarrow| + \sum_{i=d-m_-+1}^d |\gamma_+ - r_i^\downarrow| = \varepsilon + \varepsilon = 2\varepsilon; \quad (3.15)$$

so for any  $\varepsilon \in [0, 1]$ , we have  $r_\varepsilon^* \in B_\varepsilon^+(r)$ .

To summarize, we construct  $r_\varepsilon^*$  as follows: we decrease the  $m_+$  largest entries of  $r$  by setting them to  $\gamma_+$  (where  $m_+$  and  $\gamma_+$  are related by eq. (3.10)), increase the  $m_-$  smallest entries of  $r$  by setting them to  $\gamma_-$  (where  $m_-$  and  $\gamma_-$  are related by eq. (3.12)), and we keep the other entries of  $r$  unchanged. This is illustrated in Figure 3.4, for a state  $r \in \mathcal{D}(\mathcal{H})$  with  $\varepsilon = 0.07$  and  $d = 12$ .

In the case of a quantum state  $\rho \in \mathcal{D}(\mathcal{H})$ , let the sorted eigendecomposition of  $\rho$  be given as  $\rho = \sum_{i=1}^d r_i^\downarrow U |i\rangle\langle i| U^*$ , so  $(r_i^\downarrow)_{i=1}^d \in \mathcal{P}^\downarrow$  is the sorted collection of eigenvalues of  $\rho$ , counted with multiplicity (note  $\langle i| = e_i^T$  can be seen as a row vector). Then simply define

$$\rho_\varepsilon^* := \sum_{i=1}^d (r_\varepsilon^*)_i^\downarrow U |i\rangle\langle i| U^*. \quad (3.16)$$

This choice ensures that  $\rho_\varepsilon^* \in B_\varepsilon(\rho)$  and

$$(r^\downarrow)_\varepsilon^* \prec q \quad \forall q \in B_\varepsilon(r^\downarrow) \quad \implies \quad \rho_\varepsilon^* \prec \omega \quad \forall \omega \in B_\varepsilon(\rho) \quad (3.17)$$

as discussed in Section 3.1.

In the next section, we provide three ways to motivate this construction. Then in Section 3.5, we prove that  $r_\varepsilon^*$  satisfies the majorization relation (3.7).

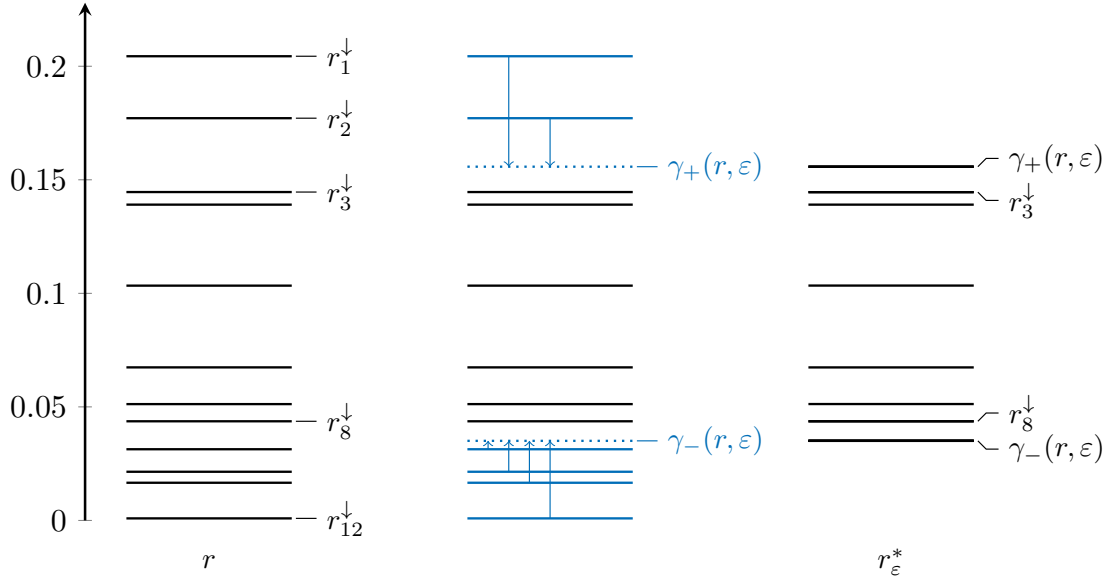


Fig. 3.4 We choose  $d = 12$ , a probability vector  $r \in \mathcal{P}_d$ , and  $\varepsilon = 0.07$ , for which  $m_+ = 2$  and  $m_- = 4$ . Left: the entries  $r_1^\downarrow \geq r_2^\downarrow \geq \dots \geq r_d^\downarrow$  of  $r$  are plotted. Center: the smallest four entries of  $r$  are increased to  $\gamma_- = \frac{1}{4}[r_d^\downarrow + \dots + r_{d+4}^\downarrow + \varepsilon]$ , and the largest two entries of  $r$  decreased to  $\gamma_+ = \frac{1}{2}[r_1^\downarrow + r_2^\downarrow - \varepsilon]$ . Right: the entries of  $r_\varepsilon^*$  are  $\gamma_+$  with multiplicity two,  $r_3^\downarrow, r_4^\downarrow, \dots, r_{d-4}^\downarrow$ , and  $\gamma_-$  with multiplicity four.

### 3.4 Three ways to find the majorization-minimizer

In this section, we present three ways to derive the form of the state given in (3.13).

#### 3.4.1 From optimality conditions

One way to obtain  $\rho_\varepsilon^*$  is from the following theorem, which is a condensed form of Theorem 7.3.1 in Chapter 7. Given a suitable function  $\varphi : \mathcal{D}(\mathcal{H}) \rightarrow \mathbb{R}$ , and any state  $\rho \in \mathcal{D}(\mathcal{H})$ , the theorem provides a necessary and sufficient condition under which a state maximizes  $\varphi$  in the  $\varepsilon$ -ball (of positive definite states),  $B_\varepsilon^+(\rho)$ , of the state  $\rho$ . These are simply the first-order optimality conditions provided by Fermat's theorem; see Chapter 7 for more details.

**Theorem 3.4.1.** *Let  $\rho \in \mathcal{D}(\mathcal{H})$ ,  $\varepsilon > 0$ , and  $\varphi : \mathcal{D}(\mathcal{H}) \rightarrow \mathbb{R}$  be a concave, continuous function which is Gâteaux-differentiable<sup>2</sup> on  $\mathcal{D}_+(\mathcal{H})$ . A state  $\xi \in B_\varepsilon^+(\rho)$ , satisfies  $\xi \in \operatorname{argmax}_{B_\varepsilon(\rho)} \varphi$  if and only if both of the following conditions are satisfied. Here  $L_\xi := \nabla \varphi(\xi)$  denotes the Gâteaux-gradient of  $\varphi$  at  $\xi$ .*

<sup>2</sup>For the definition of Gateaux-differentiability and Gateaux gradient see Section 7.1.

1. Either  $\frac{1}{2}\|\xi - \rho\|_1 = \varepsilon$  or  $L_\xi = \lambda \mathbf{1}$  for some  $\lambda \in \mathbb{R}$ , and
2. we have

$$\pi_\pm L_\xi \pi_\pm = \lambda_\pm (L_\xi) \pi_\pm, \quad (3.18)$$

where  $\pi_\pm$  is the projection onto the support of  $\Delta_\pm$ , and where  $\Delta = \Delta_+ - \Delta_-$  is the Jordan decomposition of  $\Delta := \xi - \rho$ , i.e.  $\Delta_\pm \geq 0$ .

As shown in Chapter 7, the von Neumann entropy  $S$  satisfies the requirements of the function  $\varphi$  of the theorem, with  $L_\xi = -\log \xi - \frac{1}{\log_e(2)} \mathbf{1}$ . Using Theorem 3.4.1, we deduce properties of and the form of a maximizer of  $S$  in the  $\varepsilon$ -ball.

Since  $S$  is continuous and  $B_\varepsilon(\rho)$  is compact,  $S$  achieves a maximum over  $B_\varepsilon(\rho)$ . Moreover, since  $S$  is strictly concave, the maximum is unique; otherwise, if  $\xi_1, \xi_2 \in B_\varepsilon(\rho)$  were maximizers,  $\xi = \frac{1}{2}\xi_1 + \frac{1}{2}\xi_2 \in B_\varepsilon(\rho)$  would have strictly higher entropy.

Now, let  $\xi$  be the unique maximizer. Condition 1 of Theorem 3.4.1 yields that either  $\log \xi \propto \mathbf{1}$ , so  $\xi = \tau := \frac{1}{d}$ , or else  $T(\xi, \rho) = \epsilon$ . Since  $\tau$  is the global maximizer of  $S$  over  $\mathcal{D}$ , we have  $\xi = \tau$  whenever  $\tau \in B_\varepsilon(\rho)$ . If  $\tau \notin B_\varepsilon(\rho)$ , then this condition yields the first piece of information about the maximizer: it is on the boundary of  $B_\varepsilon(\rho)$ , in that  $T(\xi, \rho) = \epsilon$ , just as shown in Figure 3.3b in the classical setup.

By (3.4), working in the basis in which  $\rho = \text{Eig}^\downarrow(\rho)$ , we have

$$T(\text{Eig}^\downarrow(\xi), \rho) \leq T(\xi, \rho) \leq \epsilon$$

and therefore  $\text{Eig}^\downarrow(\xi) \in B_\varepsilon(\rho)$ . Since  $S$  is unitarily invariant,  $S(\text{Eig}^\downarrow(\xi)) = S(\xi)$ , and by uniqueness of the maximizer, we have  $\xi = \text{Eig}^\downarrow(\xi)$ . Hence, the maximizer  $\xi$  commutes with  $\rho$ , and hence with  $\Delta$ , and the sums of its eigenprojections  $\pi_\pm$ . Since  $[L_\xi, \xi] = 0$  as well, Theorem 3.4.1 yields

$$L_\xi \pi_\pm = \lambda_\pm (L_\xi) \pi_\pm \quad (3.19)$$

For any  $\psi \in \pi_\pm \mathcal{H}$ , we have  $L_\xi \psi = \lambda_\pm (L_\xi) \psi$ , so (3.19) is an eigenvalue equation for  $L_\xi$ . Since  $\xi = \exp(-(L_\xi + \mathbf{1}))$  is a function of  $L_\xi$ , it shares the same eigenprojections. In particular,

$$\xi \pi_\pm = \exp(-(\lambda_\pm (L_\xi) + 1)) \pi_\pm$$

serves as an eigenvalue equation for  $\xi$ . Since  $[\xi, \rho] = 0$ , we can discuss how each acts on each (shared) eigenspace. By definition,  $\xi$  and  $\rho$  act the same on  $\ker \Delta = \ker(\xi - \rho)$ . On the other hand, on the subspaces where the eigenvalues of  $\xi$  are greater than those of  $\rho$ , i.e. on  $\pi_+ \mathcal{H}$ , we see that  $\xi$  has the constant eigenvalue  $\gamma_- := \exp(-(\lambda_+(L_\xi) + 1))$ ,

and on  $\pi_- \mathcal{H}$ ,  $\gamma_+ := \exp(-(\lambda_-(L_\xi) + 1))$ . Note that since  $x \mapsto -\log x - 1$  is monotone decreasing on  $\mathbb{R}$ , the subspace  $\pi_+ \mathcal{H}$  where  $L_\xi$  has its largest eigenvalue,  $\xi$  has its smallest eigenvalue, and vice-versa. That is,  $\lambda_+(\xi) = \gamma_+$  and occurs on the subspace  $\pi_- \mathcal{H}$ , and  $\lambda_-(\xi) = \gamma_-$ , and occurs on  $\pi_+ \mathcal{H}$ . Let us briefly remark on the notation:  $\gamma_-$ , for example, has the subscript  $-$  as  $\gamma_-$  is the smallest eigenvalue of  $\xi$ , whereas  $\pi_+$  is the projector onto the subspace on which eigenvalues of  $\rho$  *increase* to form the eigenvalues of  $\xi$  (which are, in fact,  $\gamma_-$ ).

Let us summarize the above observations. In  $\ker \Delta$ , the maximizer  $\xi$  has the same eigenvalues as  $\rho$ . On the subspace  $\pi_- \mathcal{H}$ , the state  $\xi$  has the constant eigenvalue  $\gamma_+$ , which is the largest eigenvalue of  $\xi$ , and  $\xi \pi_- = \gamma_+ \pi_- \leq \rho \pi_-$ . In the subspace  $\pi_+ \mathcal{H}$ ,  $\xi$  has the constant eigenvalue  $\gamma_-$ , which is its smallest eigenvalue, and  $\xi \pi_+ = \gamma_- \pi_+ \geq \rho \pi_+$ . It remains to choose subspaces corresponding to  $\pi_\pm$ . The associated eigenvalues  $\gamma_-$  and  $\gamma_+$  are then determined by  $\text{tr}[(\xi - \rho)_+] = \text{tr}[(\xi - \rho)_-] = \epsilon$ , using that  $T(\xi, \rho) = \epsilon$ .

These values, of course, turn out to be the same  $\gamma_\pm$  from Section 3.3. How might one guess  $\pi_\pm$ , however? As the entropy is minimized on pure states and maximized on the completely mixed state  $\tau$ , one can guess that to increase the entropy, one should raise the small eigenvalues of  $\rho$ , and lower the large eigenvalues of  $\rho$ ; moreover, from the non-linearity of  $-x \log x$ , one might guess to maximize the entropy it helps more to raise the *smallest* eigenvalues of  $\rho$  and lower the *largest* eigenvalues. That is,  $\pi_+$  should correspond to the eigenspaces of the  $m_-$  smallest eigenvalues of  $\rho$ , and  $\pi_-$  should correspond to the eigenspaces of the  $m_+$  largest eigenvalues of  $\rho$ , for some  $m_-, m_+ \in \{1, \dots, d-1\}$ . This turns out to be correct, as described in Section 3.3. Moreover, as  $\gamma_+$  is the largest eigenvalue of  $\xi$ , and  $\gamma_-$  is the smallest one, we must have  $\gamma_- \leq \mu \leq \gamma_+$  for any eigenvalue  $\mu$  of  $\rho$  with corresponding eigenspace which is a subspace of  $\ker \Delta$ . In Lemma 3.5.2, we prove there exists unique  $\gamma_-, \gamma_+, m_-$  and  $m_+$  which respect these considerations. Following the notation of Section 3.3, we call the resulting state  $\rho_\epsilon^*$  (instead of  $\xi$ ).

### 3.4.2 Waterfilling argument

The discussion in Section 3.4.1 uses the tools of Chapter 7 which are applicable to any differentiable concave function over  $B_\epsilon(\rho)$ . However, in the specific case of maximizing the Shannon entropy over the total-variation distance ball, a simpler argument known as water-filling suffices to motivate the form of (3.13). For the sake of brevity, we will only sketch the original argument and its applicability to motivating the form of (3.13) here, rather than including all the details.

As described by Boyd and Vandenberghe in [BV04, Example 5.2], this argument goes as follows. One considers  $n$  classical communication channels, each with capacity

$C(\vec{x}) = -\log(c_i + x_i)$  for  $i = 1, \dots, n$  with constants  $c_i > 0$ , and variables  $x_i \geq 0$  representing the percentage of power allocated to channel  $i$ , subject to the constraint  $\sum_{i=1}^n x_i = 1$ . By checking the Karush-Kuhn-Tucker (KKT) conditions from the theory of differentiable convex optimization, one finds that  $C(\vec{x})$  is maximized by  $\vec{x}^* = (x_i^*)$  where  $x_i^* = \max\{0, z^* - c_i\}$  with  $z^*$  uniquely determined by the constraint  $\sum_{i=1}^n x_i^* = 1$ . This solution has the interpretation of waterfilling: one imagines the ground composed of  $n$  patches, with the ground at height  $c_i$  over patch  $i$ . One floods the ground to a depth  $z^*$ , using the total amount of water  $\sum_{i=1}^n \max\{0, z^* - c_i\}$ , and continues pouring water until a total amount 1 of water is used. Then the amount of water added to patch  $i$  is the optimal value  $x_i^*$ .

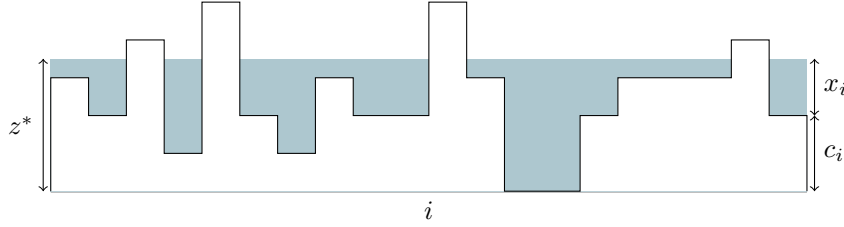


Fig. 3.5 Depiction of the waterfilling solution. The amount of water at site  $i$  is given by  $x_i^* = \max\{0, z^* - c_i\}$  where the depth  $z^*$  is determined by the constraint  $\sum_{i=1}^n x_i^* = 1$ . Based on [BV04, Figure 5.7].

In fact, one finds the same solution if the objective function is changed to  $\tilde{C}(\vec{x}) = \sum_{i=1}^n -(c_i + x_i) \log(c_i + x_i)$  subject to the constraint  $\sum_{i=1}^n x_i = \varepsilon$ , namely  $x_i^* = \max\{0, z^* - c_i\}$  with  $z^*$  determined by  $\sum_{i=1}^n x_i^* = \varepsilon$ .

Now, let us return to optimizing the Shannon entropy over the  $\varepsilon$ -ball, and set the unique optimizer as  $q \in B_\varepsilon(p)$ . If we fix  $I_- \subseteq \{1, \dots, n\}$  as the set of indices such that  $p_i \geq q_i$ , we have  $\sum_{i \in I_-} (p_i - q_i) = \varepsilon$ , and determining  $x_i := p_i - q_i$  with  $i \in I_-$  is exactly the problem of optimizing  $\tilde{C}$  with the constraints  $x_i \geq 0$ ,  $\sum_{i \in I_-} x_i = \varepsilon$ , yielding a constant “water depth”  $z_-$ . By following the analogous argument with the set  $I_+ = \{i : p_i < q_i\}$ , yielding water depth  $z_+$ , one may see that  $p_i \in \{q_i, z_-, z_+\}$ , depending on whether  $p_i > q_i$ ,  $p_i = q_i$ , or  $p_i < q_i$ . The difficulty, then, is in determining the index sets  $I_\pm$ . This is exactly the task of determining the subspaces  $\pi_\pm$  discussed in the previous section.

### 3.4.3 From majorization-infimum

A third way to establish to find  $r_\varepsilon^*$  is via the majorization-infimum. This method has the advantage of providing an algorithm to exactly calculate  $r_\varepsilon^*$ , although it's not



immediate to derive the form (3.13) from it<sup>3</sup>. Beyond providing a means to construct  $r_\varepsilon^*$ , by investigating the connection to the majorization-infimum, we find more insight into an apparent discrepancy between the majorization-minimizer and the majorization-maximizer: namely, that the majorization-maximizer of  $B_\varepsilon(p)$  lies on a vertex of the polytope, while the majorization-minimizer in general does not (see Section 3.2). As we will see in Proposition 3.4.5, the majorization-minimizer does lie on a vertex of one of a particular family of polytopes which make up  $B_\varepsilon(p)$ .

Majorization, as defined in (2.3), is a pre-order, meaning it is reflexive ( $p \prec p$ ) and transitive ( $p \prec q \prec r \implies p \prec r$ ), and on the set of sorted probability vectors,

$$\mathcal{P}^\downarrow = \left\{ p \in \mathbb{R}^d : p_1 \geq \dots \geq p_d \geq 0, \sum_{i=1}^d p_i = 1 \right\},$$

majorization is a partial order, meaning it is also antisymmetric ( $p \prec q$  and  $q \prec p$  implies  $p = q$ ). It also satisfies the so-called *lattice* property [CV02], meaning for any pair  $p, q \in \mathcal{P}^\downarrow$ , there is a unique greatest lower bound  $\inf_{\prec}(p, q) \in \mathcal{P}^\downarrow$ , which satisfies

- $\inf_{\prec}(p, q) \prec p$  and  $\inf_{\prec}(p, q) \prec q$
- if  $r \in \mathcal{P}^\downarrow$  is any other lower bound, meaning  $r \prec p$  and  $r \prec q$ , then  $r \prec \inf_{\prec}(p, q)$ .

Likewise, there is a unique least upper bound  $\sup_{\prec}(p, q) \in \mathcal{P}^\downarrow$  such that  $p \prec \sup_{\prec}(p, q)$ ,  $q \prec \sup_{\prec}(p, q)$ , and  $\sup_{\prec}(p, q) \prec r$  for any  $r \in \mathcal{P}^\downarrow$  satisfying the relations  $p \prec r$  and  $q \prec r$ . [CV02] provides an explicit algorithms to construct  $\inf_{\prec}(p, q)$  and  $\sup_{\prec}(p, q)$ .

In fact, majorization satisfies the stronger *complete lattice* property, meaning the infimum and supremum of an arbitrary subset  $\mathcal{S}$  of  $\mathcal{P}^\downarrow$  exist and are unique, and can be obtained by an explicit algorithm [YG19] (see also [BBHF+19]). For a possibly non-sorted set  $\mathcal{S} \subseteq \mathcal{P}$ , the supremum and infimum exist but are non-unique, since any permutation of a supremum (resp. infimum) is another supremum (resp. infimum). In this case, we will let  $\inf_{\prec} \mathcal{S}$  be the collection of majorization infima of  $\mathcal{S}$ , and likewise  $\sup_{\prec} \mathcal{S}$  be the collection of majorization suprema of  $\mathcal{S}$ .

As shown by [YG19] and [BBHF+19], in the case of a general set  $\mathcal{S} \subseteq \mathcal{P}$ , one can construct a majorization-infimum (which lies in  $\mathcal{P}$  but possibly not  $\mathcal{S}$ ) as follows. Let

$$s_k := y_k - y_{k-1} \tag{3.20}$$

---

<sup>3</sup>In the case that  $\varepsilon$  is small (i.e.  $\varepsilon < \delta(r)$ , in the notation of Section 4.3.1), it is straightforward, however.

where  $y_0 := 0$  and  $y_k := \inf_{q \in \mathcal{S}} \sum_{i=1}^k q_i^\downarrow$ . Then  $s \in \inf_{\prec} \mathcal{S}$ . Note that  $s \in \mathcal{P}^\downarrow$  by construction.

Since Theorem 3.2.1 shows that  $r_\varepsilon^* \in \inf_{\prec} B_\varepsilon(r)$ , equation (3.20) provides another means to construct it. However, Theorem 3.2.1 shows something stronger: not only is  $r_\varepsilon^* \in \inf_{\prec} B_\varepsilon(r)$ , but  $r_\varepsilon^* \in B_\varepsilon(r)$  too (it is a minimum, not just an infimum).

Note that for a convex set  $\mathcal{S} \subseteq \mathcal{P}$ , if it has a majorization-minimum, then that minimum is unique. To see this, simply note that there exist functions  $\mathcal{P} \rightarrow \mathbb{R}$  which are both strictly concave and Schur concave, such as the Shannon entropy  $S$ . If  $r_1, r_2 \in \mathcal{S}$  were distinct majorization-minima, then  $S$  would attain its maximum over  $\mathcal{S}$  at each of them, by Schur concavity. However, if  $\mathcal{S}$  is convex, then  $r := \frac{1}{2}r_1 + \frac{1}{2}r_2 \in \mathcal{S}$  too, and by strict convexity, then  $S(r) > \frac{1}{2}S(r_1) + \frac{1}{2}S(r_2)$ , a contradiction.

The following proposition shows that (3.20) can be used to establish a necessary and sufficient condition for the majorization-infimum to be a minimum.

**Proposition 3.4.2.** *Let  $\mathcal{S} \subseteq \mathcal{P}$  be an arbitrary set of probability vectors.  $\mathcal{S}$  admits a majorization minimum if and only if for some  $q^* \in \mathcal{S}$  the infimum*

$$\inf_{q \in \mathcal{S}} \sum_{i=1}^k q_i^\downarrow \quad (3.21)$$

*is achieved at  $q^*$  for all  $k = 1, 2, \dots, d$ . In the latter case,  $q^*$  is a majorization-minimum of  $\mathcal{S}$ .*

*Proof.* In the following, we use the notation of (3.20) and that  $s \in \mathcal{P}^\downarrow$ , defined in that equation, is a majorization-infimum of  $\mathcal{S}$ . If the infimum in (3.21) is achieved at  $q^*$  for each  $k = 1, \dots, d$ , then

$$s_k = y_k - y_{k-1} = \sum_{i=1}^k (q_i^*)^\downarrow - \sum_{i=1}^{k-1} (q_i^*)^\downarrow = (q_k^*)^\downarrow$$

and hence  $s = (q^*)^\downarrow$ . As a permutation of  $s$ , therefore,  $q^* \in \inf_{\prec} \mathcal{S}$ . Since  $q^* \in \mathcal{S}$ ,  $q^*$  is a majorization minimum of  $\mathcal{S}$ .

On the other hand, assume that  $\mathcal{S}$  admits a majorization minimum  $q^* \in \mathcal{S}$ . Since  $q^* \in \inf_{\prec} \mathcal{S}$ ,  $q^* \prec q$  for all  $q \in \mathcal{S}$ . By definition, this means that

$$\sum_{i=1}^k (q_i^*)^\downarrow \leq \sum_{i=1}^k q_i^\downarrow \quad \forall q \in \mathcal{S}$$

and hence infimum in (3.21) is achieved on  $q^*$  for each  $k = 1, \dots, d$ .  $\square$

*Remark.* Proposition 3.4.2 is not immediately helpful for establishing computationally that a given set  $\mathcal{S}$  admits no majorization-minima, because the optimization in (3.21) could have multiple solutions. Hence, solving the optimization problem in (3.21) for two different choices of  $k$  and obtaining two different optima is not sufficient to prove that  $\mathcal{S}$  does not admit a majorization-minimum. Instead, in Section 3.7, to establish that balls in  $p$ -norm with  $1 < p < \infty$  do not admit a majorization-minimizer, the majorization-infimum is constructed via (3.20) and shown to lie outside of the ball itself.

Before moving to the following section, in which we prove the construction of  $r_\varepsilon^*$  given in Section 3.3 truly yields the majorization minimum, let us take brief detour to investigate the majorization-infimum further.

First, a word on the computational aspect of calculating the  $y_k$  in (3.20). Since the objective  $\sum_{i=1}^k q_i^\downarrow$  in the optimization in  $y_k$  is semidefinite representable, when membership in  $\mathcal{S}$  can be described by linear or semidefinite constraints, the optimization in  $y_k$  can be represented as a semidefinite program. In fact, when  $\mathcal{S} \subseteq \mathcal{P}$  can be described by a set of linear constraints in the sense that

$$\mathcal{S} = \{q : Aq \geq \alpha, Bq = \beta\}$$

for some matrices  $A$  and  $B$  and vectors  $\alpha$  and  $\beta$ , [YG19] shows that the majorization-infimum can be computed by a linear program. One way to see<sup>4</sup> the construction of  $r_\varepsilon^*$  in Section 3.3 then is as providing a closed-form solution for the majorization-infimum in the special case in which  $\mathcal{S} = B_\varepsilon(r)$ .

**The majorization-minimizer of polytopes** As discussed in [BBHF+19, Lemma 1], for a polytope  $Q \subseteq \mathcal{P}^\downarrow$ , the optimization in the definition of  $y_k$  (see (3.20)) may be restricted to the vertices of the polytope; this follows from the fact that when  $Q \subseteq \mathcal{P}^\downarrow$ ,  $y_k$  constitutes a linear program. Since  $B_\varepsilon(r)$  is a polytope (recalling the discussion in Section 3.2), the case of polytopes is quite relevant to us.

There are two apparent difficulties, however, to applying this result to the majorization-minimizer of polytopes. First, Proposition 3.4.2 shows that the existence of a majorization-minimizer is equivalent to the optimization in the definition of  $y_k$  being achieved on the same element for each  $k$ , and as noted in the following remark, the optimization in the  $y_k$  may obtain multiple solutions. In that case, while one of the solutions may be a vertex of the polytope, the common solution, a priori, could lie in the interior. This is

---

<sup>4</sup>This point of view was shared with me by Xiao-Dong Yu in private communication.

ruled out (in the more general case of the union of polytopes) by the following result, Proposition 3.4.4.

**Lemma 3.4.3.** *Let  $V \subseteq \mathcal{P}^\downarrow$  be a finite set of sorted probability vectors. Let  $\text{conv } V := \{\sum_i \lambda_i v_i : \lambda_i \geq 0, \sum_i \lambda_i = 1, v_i \in V\}$  be its convex hull. Then*

$$p \prec V \iff p \prec \text{conv}(V) \quad (3.22)$$

where we use the notation  $p \prec \mathcal{S}$  for a set  $\mathcal{S} \subseteq \mathcal{P}$  to mean  $p \prec s$  for each  $s \in \mathcal{S}$ . Moreover if  $p \prec V$  and  $p \in \text{conv } V$  then  $p \in V$ .

*Proof.* The implication “ $\Leftarrow$ ” of (3.22) is trivial since  $V \subseteq \text{conv } V$ . On the other hand, assume  $p \prec V$  and  $q \in \text{conv } V$  is given by  $q = \sum_j \lambda_j v^{(j)}$  for a probability distribution  $\{\lambda_i\}_i$  and  $v^{(j)} \in V$ . Since for each  $k$  and  $j$ ,

$$\sum_{i=1}^k p_i \leq \sum_{i=1}^k v_i^{(j)}$$

we have

$$\sum_{i=1}^k p_i \leq \sum_j \lambda_j \sum_{i=1}^k v_i^{(j)} = \sum_{i=1}^k q_i.$$

As each vector lies in  $\mathcal{P}^\downarrow$ , we conclude that  $p \prec q$ .

Next, assume  $p \prec V$  and  $p \in \text{conv } V$  is given by  $p = \sum_j \lambda_j v^{(j)}$  for a probability distribution  $\{\lambda_i\}_i$  and  $v^{(j)} \in V$ . Then let  $S_k^{(j)} = \sum_{i=1}^k v_i^{(j)}$ . Set  $J := \{j : \lambda_j > 0\}$ . We have that for each  $k = 1, \dots, d$ ,

$$\sum_{i=1}^k q_i = \sum_{i=1}^k \sum_{j \in J} \lambda_j v_i^{(j)} = \sum_{j \in J} \lambda_j S_k^{(j)} \leq \sum_{i=1}^k v_i^{(j)} = S_k^{(j)} \text{ for all } j \in J.$$

Hence,  $\sum_{j \in J} \lambda_j S_k^{(j)} \leq \min_{j \in J} S_k^{(j)}$ . For the average to be smaller than the minimum, the distribution must be constant; in other words, we must have  $S_k^{(j)} \equiv c_k := \min_{j \in J} S_k^{(j)}$  for each  $j \in J$ . Thus, for  $j \in J$ ,

$$\sum_{i=1}^k v_i^{(j)} = c_k \implies v_k^{(j)} = c_k - c_{k-1} \quad \forall k$$

where we set  $c_0 = 0$ . That is, there can only be one vector  $v^{(j)}$  such that  $j \in J$ , since  $j \in J$  completely fixes the elements of  $v^{(j)}$ . Hence,  $q = v^{(j)}$  for some  $j$ , and therefore  $q \in V$  as desired.  $\square$

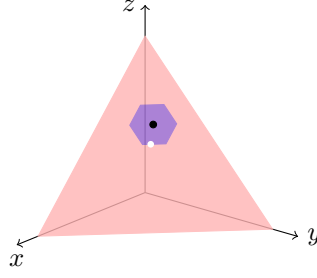


Fig. 3.6 The example of Figure 3.3, again: In dimension  $d = 3$ , the simplex,  $\Delta$ , of probability vectors is the shaded triangle, along with the ball  $B_\epsilon(p)$  which is the hexagon shown in blue, centered at  $p = (0.21, 0.24, 0.55)$  (depicted by a black dot) with  $\epsilon = 0.1$ . The majorization-minimizer  $p_\epsilon^*$  is drawn in white.

This lemma can be used to show that majorization-minimizers lie on vertices, as follows.

**Proposition 3.4.4.** *Let  $Q \subseteq \mathcal{P}^\downarrow$  be a union of polytopes,  $Q = \bigcup_i Q_i$ . Let  $V(Q_i)$  be the vertices of  $Q_i$ , and  $V(Q) = \bigcup_i V(Q_i)$ . Then  $q^* \in \mathcal{P}^\downarrow$  is a majorization-minimizer of  $V(Q)$  if and only if  $q^*$  is a majorization-minimizer of  $Q$ .*

*Proof.* If  $V(Q)$  admits a majorization-minimizer  $q^*$ , then  $q^* \prec V(Q)$ . In particular, for each  $i$ ,  $q^* \prec V(Q_i)$  and hence  $q^* \prec Q_i$  by the lemma. Thus,  $q^* \prec Q$ ; since  $q^* \in V(Q) \subseteq Q$ ,  $q^*$  is a majorization-minimizer of  $Q$ .

On the other hand, if  $Q$  admits a majorization-minimizer  $q^*$ , then  $q^* \in Q_i$  for some  $i$ . Hence  $q^* \prec V(Q_i)$  and  $q^* \in \text{conv } V(Q_i)$ , so by the lemma,  $q^* \in V(Q_i) \subseteq V(Q)$ . Since  $q^* \prec V(Q)$ ,  $q^*$  is a majorization minimizer of  $V(Q)$ .  $\square$

The second difficulty is that for a polytope  $Q \subseteq \mathcal{P}$  which may not be sorted, the majorization-minimum may not lie on a vertex of  $Q$ . For example, see Figure 3.6 for an example in which the majorization-minimizer  $p_\epsilon^*$  does not lie on a vertex of  $B_\epsilon(p)$ .

However, if  $\mathcal{S} \subseteq \mathcal{P}$  is a polytope, then  $\mathcal{S}^\downarrow$  can be seen as a union of polytopes, as follows. For a permutation  $\pi \in S_d$ , let

$$E_\pi := \{q \in \mathcal{P} : q_{\pi(1)} \geq q_{\pi(2)} \geq \cdots \geq q_{\pi(d)}\} \quad (3.23)$$

be the part of the simplex  $\mathcal{P}$  whose elements are sorted according to  $\pi$ . Note  $\mathcal{P}^\downarrow = E_{(1,2,3)}$ . As  $E_\pi$  can be written as  $\mathcal{P}$  intersected with  $d-1$  halfspaces, with each halfspace enforcing an inequality  $q_{\pi(i)} \geq q_{\pi(i+1)}$  for  $i = 1, \dots, d-1$ , we have that  $E_\pi$  is a polytope. The sets  $E_\pi$  are shown for  $d = 3$  in Figure 3.7.

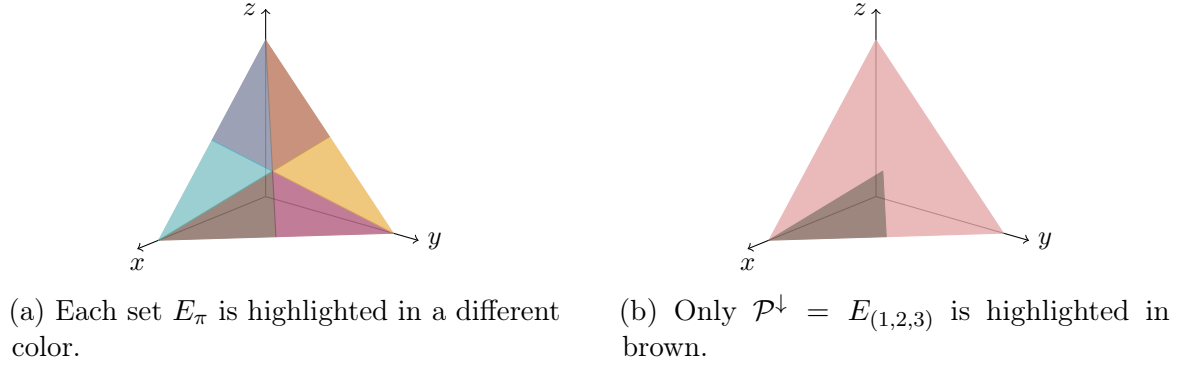


Fig. 3.7 The simplex probability  $\mathcal{P} \subseteq \mathbb{R}^3$  with  $d = 3$ , decomposed into the six components  $\mathcal{P} = \bigcup_{\pi \in S_3} E_\pi$ .

We have that  $\bigcup_{\pi \in S_d} E_\pi = \mathcal{P}$ . Moreover, for any polytope  $\mathcal{S}$  and  $\pi \in S_d$ , the intersection  $\mathcal{S}_\pi := \mathcal{S} \cap E_\pi$  is a polytope. We can see  $B_\varepsilon(q)$  decomposed in this way in Figure 3.8, using the same example as from Section 3.2.

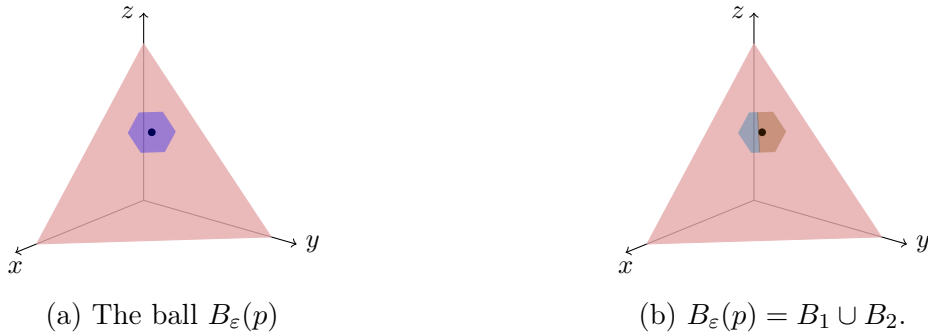


Fig. 3.8 Continuing the example of Figure 3.6, we depict  $B_\varepsilon(p)$  decomposed according to the  $\{E_\pi\}_{\pi \in S_d}$ . In this example, only two of the components are non-zero,  $B_1 := E_{(3,1,2)} \cap B_\varepsilon(p)$  and  $B_2 := E_{(3,2,1)} \cap B_\varepsilon(p)$ .

In fact,  $\mathcal{S}_\pi^\downarrow$  is a polytope too, as it is just a rotation of  $\mathcal{S}_\pi$  (in contrast to  $\mathcal{S}^\downarrow$ , in general). In particular, if  $\mathcal{S}_\pi = \text{conv}\{q_1, \dots, q_n\}$ , then  $\mathcal{S}_\pi^\downarrow = \text{conv}\{q_1^\downarrow, \dots, q_n^\downarrow\}$ . So we can obtain  $\mathcal{S}^\downarrow = \bigcup_{\pi \in S_d} \mathcal{S}_\pi^\downarrow$  in this way. This is illustrated in Figure 3.9.

Lastly, let us return to the discussion of the majorization minimizer.

**Proposition 3.4.5.** *If  $Q \subseteq \mathcal{P}$  is a polytope (such as a total-variation ball) and admits a majorization-minimizer, then it lies on one of the vertices of the polytope  $Q \cap E_\pi$  for some  $\pi \in S_d$ , where  $E_\pi$  is defined in (3.23).*

*Proof.* As discussed in the text above,  $Q^\downarrow = \bigcup_{\pi} (Q \cap E_\pi)^\downarrow$ . Since  $Q$  admits a majorization-minimizer,  $Q^\downarrow$  does too. Hence Proposition 3.4.4 shows that there is a majorization-

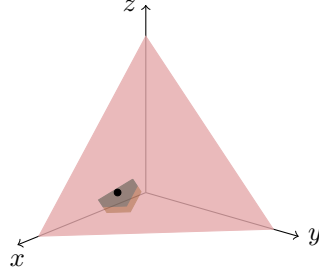


Fig. 3.9 Using the notation of Figure 3.8, we find the sorted set  $(B_\varepsilon(p))^\downarrow = B_1^\downarrow \cup B_2^\downarrow$  by sorting each components  $B_1$  and  $B_2$ .

minimizer of  $Q^\downarrow$  such that  $(q^*)^\downarrow \in V((Q \cap E_\pi)^\downarrow)$  for some  $\pi$ . Then there is a permutation  $q^*$  of  $(q^*)^\downarrow$  that lies in  $Q \cap E_\pi$ , and  $q^*$  is a majorization-minimizer of  $Q$ .  $\square$

We can see Proposition 3.4.5 illustrated in our running example in Figure 3.10.



(a) Figure 3.9 with  $(p_\varepsilon^*)^\downarrow$  marked in white. (b) Figure 3.8b with  $(p_\varepsilon^*)^\downarrow$  marked in white.

Fig. 3.10 Continuing the example of Figure 3.6, we see that in agreement with Proposition 3.4.4,  $(B_\varepsilon(p))^\downarrow$  admits its majorization-minimizer on a vertex. Returning to the original, unsorted, form in Figure 3.10b, we see that indeed,  $p_\varepsilon^*$  lies on a vertex of  $B_\varepsilon(p) \cap E_\pi$  for some  $\pi$  (in this case, either  $\pi = (3, 1, 2)$  or  $\pi = (3, 2, 1)$ ).

*Remark.* In the case of the particular polytope given by the total-variation ball  $B_\varepsilon(r)$ , one in fact does not need to consider a union of polytopes, since (3.4) shows that if  $r \in E_\pi$  for some  $\pi$ , then for any  $q \in B_\varepsilon(r)$ , for some permutation  $\pi' \in S_d$ , we have  $\pi'(q) \in E_\pi$ . In particular, we can take  $r_\varepsilon^* \in B_\varepsilon(r) \cap E_\pi$ . Then  $(B_\varepsilon(r) \cap E_\pi)^\downarrow$  itself is a sorted polytope, for which Proposition 3.4.4 yields that the majorization-minimizer lies on a vertex.

### 3.5 Minimality of $r_\varepsilon^*$ in the majorization order (3.7)

Given  $r \in \mathcal{P}_d^\downarrow$ . Let  $w \in B_\varepsilon(r)$ . We aim to show  $r_\varepsilon^* \prec w$ ; equivalently,  $r_\varepsilon^* \prec w^\downarrow$ . Since  $\|r^\downarrow - w^\downarrow\|_1 \leq \|r - w\|_1$ , using (3.4) (by promoting vectors to diagonal matrices), we have

$w^\downarrow \in B_\epsilon(r)$  as well. Hence, without loss of generality, we may consider  $w = w^\downarrow$ , with entries

$$w_1 \geq \cdots \geq w_d.$$

Let  $r_1 \geq \cdots \geq r_d$  be the entries of  $r$ , and  $r_1^* \geq \cdots \geq r_d^*$  be the entries of  $r_\epsilon^*$ . In the following, we show  $r_\epsilon^* \prec w$ .

1. First, we establish that  $r_1^* \leq w_1$ .

To prove this, let us assume the contrary:  $r_1^* > w_1$ . Then, since  $r_1^* = r_2^* = \cdots = r_{m_+}^* = \gamma_+$ ,

$$m_+ \gamma_+ = \sum_{i=1}^{m_+} r_i^* > m_+ w_1 \geq \sum_{i=1}^{m_+} w_i.$$

We conclude this step with the following lemma.

**Lemma 3.5.1.** *If  $m_+ \gamma_+ > \sum_{i=1}^{m_+} w_i$ , then  $w \notin B_\epsilon(r)$ .*

*Proof.* Multiplying each side by  $-1$  and adding  $\sum_{i=1}^{m_+} r_i$ , we have

$$\sum_{i=1}^{m_+} (r_i - r_i^*) < \sum_{i=1}^{m_+} (r_i - w_i) \leq \sum_{i=1}^{m_+} (r_i - w_i)_+ \leq \sum_{i=1}^d (r_i - w_i)_+ = \frac{1}{2} \|r - w\|_1. \quad (3.24)$$

Using  $\gamma_+ = \frac{1}{m_+} (\sum_{i=1}^{m_+} r_i - \epsilon)$ , the far left-hand side is

$$\sum_{i=1}^{m_+} r_i - m_+ \gamma_+ = \sum_{i=1}^{m_+} r_i - \left( \sum_{i=1}^{m_+} r_i - \epsilon \right) = \epsilon.$$

Then (3.24) becomes

$$\epsilon < \frac{1}{2} \|r - w\|_1,$$

contradicting that  $w \in B_\epsilon(r)$ . □

2. Next, for  $k \in \{1, \dots, m_+\}$ , we have  $\sum_{i=1}^k r_i^* \leq \sum_{i=1}^k w_i$ .

We prove this recursively: assume the property holds for some  $k \in \{1, \dots, m_+ - 1\}$  but not for  $k + 1$ . Note we have proven the base case of  $k = 1$  in the previous step. Then



$$\sum_{i=1}^k r_i^* \leq \sum_{i=1}^k w_i, \quad \text{and} \quad \sum_{i=1}^{k+1} r_i^* > \sum_{i=1}^{k+1} w_i. \quad (3.25)$$

Subtracting the first inequality in (3.25) from the second, we have

$$r_{k+1}^* > w_{k+1}$$

yielding  $\gamma_+ > w_{k+1} \geq w_k \geq \dots \geq w_{m_+}$ . Summing the inequalities  $r_{k+\ell}^* = \gamma_+ > w_{k+\ell}$  for  $\ell = 2, 3, \dots, m_+ - k$ , we have

$$\sum_{j=k+2}^{m_+} r_j^* > \sum_{j=k+2}^{m_+} w_j.$$

Adding this to the second inequality of (3.25), we have

$$\sum_{i=1}^{m_+} r_i^* > \sum_{i=1}^{m_+} w_i.$$

We thus conclude by Lemma 3.5.1.

3. Next, let  $k \in \{m_+ + 1, \dots, d - m_-\}$ . Assume  $\sum_{i=1}^k r_i^* > \sum_{i=1}^k w_i$ . Then

$$\sum_{i=1}^k r_i - r_i^* < \sum_{i=1}^k r_i - w_i$$

However, the left-hand side is

$$\sum_{i=1}^k r_i - r_i^* = \sum_{i=1}^{m_+} r_i - m_+ \gamma_+ = \epsilon.$$

Hence,

$$\epsilon < \sum_{i=1}^k r_i - w_i \leq \sum_{i=1}^d (r_i - w_i)_+ = \frac{1}{2} \|w - r\|_1$$

which is a contradiction.

4. Finally, we finish with a recursive proof similar to step 2. Assume the property holds for  $k \in \{d - m_-, \dots, d - 1\}$ , but not for  $k + 1$ . For this case too we have

proven the base case  $k = d - m_-$  in the previous step. We therefore assume

$$\sum_{i=1}^k r_i^* \leq \sum_{i=1}^k w_i, \quad \text{and} \quad \sum_{i=1}^{k+1} r_i^* > \sum_{i=1}^{k+1} w_i. \quad (3.26)$$

Subtracting the two equations, we have

$$r_{k+1}^* > w_{k+1}.$$

Since  $\gamma_- = r_{k+1}^*$  we have  $\gamma_- > w_{k+1} \geq w_{k+1} \geq \dots \geq w_d$ . Summing  $r_{k+\ell}^* = \gamma_- > w_{k+\ell}$  for  $\ell = 2, 3, \dots, d$ , we have

$$\sum_{i=k+2}^d r_i^* > \sum_{i=k+2}^d w_i.$$

Adding to the second inequality of (3.26), we find

$$1 = \sum_{i=1}^d r_i^* > \sum_{i=1}^d w_i$$

which contradicts the assumption that  $r \in \Delta$ . □

### 3.5.1 Uniqueness of construction for $r_\varepsilon^*$

The uniqueness claims in the construction of  $r_\varepsilon^*$  in Section 3.3 follow from the following lemma. These properties prove useful in establishing the so-called semi-group property in Proposition 4.2.4, and are not present in [HOS18; HY10].

**Lemma 3.5.2.** *Let  $r \in \mathcal{P}_d$  have sorted entries  $r^\downarrow = (r_1^\downarrow, \dots, r_d^\downarrow)$ . Let  $0 < \varepsilon < \text{TV}(r, u)$ , where  $u = (1/d, \dots, 1/d)$ , i.e.  $u \notin B_\varepsilon(r)$ . There is a unique pair  $(\gamma_-, m_-) \in [0, 1] \times \{1, \dots, d-1\}$  such that*

$$\sum_{i=d-m_-+1}^d |\gamma_- - r_i^\downarrow| = \varepsilon \quad \text{and} \quad r_{d-m_-+1}^\downarrow < \gamma_- \leq r_{d-m_-}^\downarrow, \quad (3.27)$$

and similarly, a unique pair  $(\gamma_+, m_+) \in [0, 1] \times \{1, \dots, d-1\}$  with

$$\sum_{j=1}^{m_+} |\gamma_+ - r_j^\downarrow| = \varepsilon \quad \text{and} \quad r_{m_++1}^\downarrow \leq \gamma_+ < r_{m_+}^\downarrow. \quad (3.28)$$

We define the index sets

$$I_H = \{1, \dots, m_+\}, \quad I_M = \{m_+ + 1, \dots, d - m_-\}, \quad I_L = \{d - m_- + 1, \dots, d\} \quad (3.29)$$

corresponding to the “highest”  $m_+$  entries, “middle”  $d - m_- - m_+$  entries, and “lowest”  $m_-$  entries of  $r$ , respectively.

We have the following properties. The numbers  $\gamma_-$  and  $\gamma_+$  are such that  $\gamma_- < \frac{1}{d} < \gamma_+$ , and we have the following characterizations of the pairs  $(\gamma_-, m_-)$  and  $(\gamma_+, m_+)$ : For any  $m \in \mathbb{N}$ , defining the functions

$$\gamma_-^{(m)} := \frac{1}{m} \left( \sum_{j=d-m+1}^d r_j^\downarrow + \varepsilon \right), \quad \gamma_+^{(m)} := \frac{1}{m} \left( \sum_{j=1}^m r_j^\downarrow - \varepsilon \right), \quad (3.30)$$

we have  $\gamma_- = \gamma_-^{(m_-)}$  and  $\gamma_+ = \gamma_+^{(m_+)}$ , where  $m_-$  and  $m_+$  are, respectively, the unique solutions of the following:

$$\begin{aligned} r_{d-m'_-+1}^\downarrow < \gamma_- \leq r_{d-m'_-}^\downarrow & : \quad m'_- \in \{1, \dots, d-1\}, \\ r_{m'_++1}^\downarrow \leq \gamma_+ < r_{m'_+}^\downarrow & : \quad m'_+ \in \{1, \dots, d-1\}, \end{aligned} \quad (3.31)$$

and satisfy

$$\begin{aligned} m_- &= \min\{m' \in \{1, \dots, d-\ell\} : \gamma_-^{(m')} \leq r_{d-m'}^\downarrow\}, \\ m_+ &= \min\{m' \in \{1, \dots, \ell\} : \gamma_+^{(m')} \geq r_{m'+1}^\downarrow\}, \end{aligned} \quad (3.32)$$

where  $\ell$  is defined by

$$r_{\ell+1}^\downarrow < \frac{1}{d} \leq r_\ell^\downarrow. \quad (3.33)$$

**Proof of Lemma 3.5.2** Here we prove the results pertaining to the pair  $(\gamma_+, m_+)$ . The results for the pair  $(\gamma_-, m_-)$  can be obtained analogously. Note that if any pair  $(\gamma_+, m_+) \in [0, 1] \times \{1, \dots, d-1\}$  satisfies (3.27) then we have

$$\varepsilon = \sum_{i=1}^{m_+} |\gamma_+ - r_i^\downarrow| = \sum_{i=1}^{m_+} (r_i^\downarrow - \gamma_+) = \sum_{i=1}^{m_+} r_i^\downarrow - m_+ \gamma_+,$$

implying  $\gamma_+ = \frac{1}{m_+} \left( \sum_{i=1}^{m_+} r_i^\downarrow - \varepsilon \right) = \gamma_+^{(m_+)}$ . Conversely, if for some  $m'_+ \in \{1, \dots, d-1\}$ , the corresponding value  $\gamma_+^{(m'_+)}$  satisfies  $r_{m'_++1}^\downarrow \leq \gamma_+^{(m'_+)} < r_{m'_+}^\downarrow$ , then

$$\varepsilon = \sum_{i=1}^{m'_+} r_i^\downarrow - m'_+ \gamma_+^{(m'_+)} = \sum_{i=1}^{m'_+} (r_i^\downarrow - \gamma_+^{(m'_+)}) = \sum_{i=1}^{m'_+} |\gamma_+^{(m'_+)} - r_i^\downarrow|$$

and in particular

$$\sum_{i=1}^{m'_+} |\gamma_+^{(m'_+)} - r_i^\downarrow| = \varepsilon.$$

Hence, the existence and uniqueness of  $(\gamma_+, m_+)$  satisfying (3.27) is equivalent to the existence and uniqueness of  $m_+ \in \{1, \dots, d-1\}$  such that  $\gamma_+^{(m_+)}$  satisfies  $r_{m_++1}^\downarrow \leq \gamma_+^{(m_+)} < r_{m_+}^\downarrow$ .

Next, we show that

$$m_+ = \min\{m'_+ \in \{1, \dots, \ell\} : \gamma_+^{(m'_+)} \geq r_{m'_++1}^\downarrow\}$$

by checking that the minimum exists and uniquely solves  $r_{m_++1}^\downarrow \leq \gamma_+^{(m_+)} < r_{m_+}^\downarrow$ . The proof is then completed by showing that we must have  $\gamma_+^{(m_+)} > \frac{1}{d}$ . The steps of the construction are elucidated below.

Step 1.  $\{m'_+ \in \{1, \dots, \ell\} : \gamma_+^{(m'_+)} \geq r_{m'_++1}^\downarrow\} \neq \emptyset$ .

Let us assume the contrary. Then, in particular, that  $\gamma_+^{(\ell)} < r_{\ell+1}^\downarrow$ . By substituting  $\gamma_+^{(\ell)} = \frac{1}{\ell} \sum_{i=1}^{\ell} r_i^\downarrow - \frac{1}{\ell} \varepsilon$  in this inequality, we have

$$\sum_{i=1}^{\ell} r_i^\downarrow < \ell r_{\ell+1}^\downarrow + \varepsilon \quad (3.34)$$

Using  $\sum_{i=1}^d (r_i - u_i) = 0$ , we have  $\frac{1}{2} \|r - u\|_1 = \sum_{i=1}^d \text{neg}(u_i - r_i)$ , where  $\text{neg}(x) := \min(0, x)$ . Using the definition of  $\ell$ , eq. (3.33), this can be written as

$$\frac{1}{2} \|r - u\|_1 = \sum_{i=1}^{\ell} (r_i^\downarrow - \frac{1}{d}) = \sum_{i=1}^{\ell} r_i^\downarrow - \frac{\ell}{d}. \quad (3.35)$$

Employing (3.34), and using the definition of  $\ell$ , we find that

$$\frac{1}{2}\|r - u\|_1 < \ell r_{\ell+1}^\downarrow + \varepsilon - \frac{\ell}{d} = \varepsilon + \ell(r_{\ell+1}^\downarrow - \frac{1}{d}) < \varepsilon. \quad (3.36)$$

That is,  $u \in B_\varepsilon(r)$ , which contradicts our assumption.

Step 2. The value  $x := \min\{m'_+ \in \{1, \dots, \ell\} : \gamma_+^{(m'_+)} \geq r_{m'_++1}^\downarrow\}$  solves  $r_{x+1}^\downarrow \leq \gamma_+^{(x)} < r_x^\downarrow$ .

If  $x = 1$ , then using eq. (3.30) we see that  $r_1^\downarrow > r_1^\downarrow - \varepsilon = \gamma_+^{(1)}$ , and hence  $r_{x+1}^\downarrow \leq \gamma_+^{(x)} < r_x^\downarrow$ . Otherwise, by minimality of  $x$ , we have  $\gamma_+^{(x-1)} < r_x^\downarrow$ . We first establish that for  $m'_+ \in \{2, 3, \dots, d\}$ ,

$$\gamma_+^{(m'_+-1)} < r_n^\downarrow \iff \gamma_+^{(m'_+)} < r_n^\downarrow \quad (3.37)$$

and the result follows by taking  $m'_+ = x$ . To prove (3.37), we write

$$\begin{aligned} \gamma_+^{(m'_+-1)} < r_n^\downarrow &\iff \frac{1}{m'_+-1} \left[ \sum_{j=1}^{m'_+-1} r_j^\downarrow - \varepsilon \right] < r_n^\downarrow \\ &\iff \sum_{j=1}^{m'_+-1} r_j^\downarrow - \varepsilon < (m'_+-1)r_n^\downarrow = m'_+r_n^\downarrow - r_n^\downarrow \\ &\iff \sum_{j=1}^{m'_+} r_j^\downarrow - \varepsilon < m'_+r_n^\downarrow \\ &\iff \frac{1}{m'_+} \left[ \sum_{j=1}^{m'_+} r_j^\downarrow - \varepsilon \right] < r_n^\downarrow \\ &\iff \gamma_+^{(m'_+)} < r_n^\downarrow. \end{aligned}$$

Step 3. Uniqueness of  $m_+$  satisfying  $r_{m_++1}^\downarrow \leq \gamma_+^{(m_+)} < r_{m_+}^\downarrow$ .

By substituting the inequality  $r_{m'_+-1}^\downarrow \geq r_n^\downarrow$  into (3.37), we find for  $m'_+ \in \{2, 3, \dots, d\}$ ,

$$r_n^\downarrow > \gamma_+^{(m'_+)} \implies r_{m'_+-1}^\downarrow > \gamma_+^{(m'_+-1)}. \quad (3.38)$$

Now, assume that there exists a  $y$  satisfying  $y > x > 0$  for which  $r_y^\downarrow > \gamma_+^{(y)} \geq r_{y+1}^\downarrow$ . By applying the implication (3.38) a total of  $(y - x - 1)$  times, we see that

$$r_y^\downarrow > \gamma_+^{(y)} \implies r_{y-1}^\downarrow > \gamma_+^{(y-1)} \implies \dots \implies r_{x+1}^\downarrow > \gamma_+^{(x+1)}$$

which by (3.37) is equivalent to  $\gamma_+^{(x)} < r_{x+1}^\downarrow$ . This contradicts the assumption that  $\gamma_+^{(x)} \geq r_{x+1}^\downarrow$ . Hence, such a  $y$  cannot exist.

Step 4.  $\gamma_+^{(m_+)} > \frac{1}{d}$ .

We prove this by showing that if  $\gamma_+^{(m_+)} \leq \frac{1}{d}$  then we obtain a contradiction to the assumption  $\frac{1}{2}\|r - u\|_1 > \varepsilon$  of Lemma 3.5.2.

Assume  $\gamma_+^{(m_+)} \leq \frac{1}{d}$ . Then,

$$\gamma_+^{(m_+)} = \frac{1}{m_+} \sum_{j=1}^{m_+} r_j^\downarrow - \frac{\varepsilon}{m_+} \leq \frac{1}{d} \iff \sum_{j=1}^{m_+} \left( r_j^\downarrow - \frac{1}{d} \right) \leq \varepsilon.$$

Now, since  $m_+ \leq \ell$  by (3.32),  $r_j^\downarrow - \frac{1}{d} \geq 0$  for each  $j \in \{1, \dots, m_+\}$ .

If  $m_+ = \ell$ , then

$$\sum_{j=1}^{\ell} \left( r_j^\downarrow - \frac{1}{d} \right) \leq \varepsilon. \quad (3.39)$$

On the other hand, if  $m_+ < \ell$ , then  $m_+ + 1 \leq \ell$ . Then, using the assumption  $\gamma_+^{(m_+)} \leq \frac{1}{d}$ ,

$$\frac{1}{d} \geq \gamma_+^{(m_+)} \geq r_{m_++1}^\downarrow \geq \dots \geq r_\ell^\downarrow \geq \frac{1}{d},$$

so  $r_{m_++1}^\downarrow = r_{m_++2}^\downarrow = \dots = r_\ell^\downarrow = \frac{1}{d}$ . Then, (3.39) holds in this case as well. Then by (3.35),

$$\frac{1}{2}\|r - u\|_1 = \sum_{j=1}^{\ell} \left( r_j^\downarrow - \frac{1}{d} \right) \leq \varepsilon,$$

which contradicts the assumption that  $u \notin B_\varepsilon(r)$ .  $\square$

### 3.6 Constructing the maximal state in the majorization order (3.7)

As mentioned earlier, the minimum principle tells us that the entropy is minimized on a vertex of  $B_\varepsilon(p)$ . In the example of Figure 3.2 in  $d = 3$ , with  $p = (0.21, 0.24, 0.55)$  and  $\varepsilon = 0.1$ , we considered the set  $M_w$  of points of  $\mathcal{P}$  majorized by  $w \in B_\varepsilon(p)$ . In the same example in Figure 3.3, the minimum of the entropy over  $B_\varepsilon(p)$  was found to be  $p_{*,\varepsilon} = (0.21 - \varepsilon, 0.24, 0.55 + \varepsilon)$ . In Figure 3.11, we see that in fact  $B_\varepsilon(p) \subseteq M_{p_{*,\varepsilon}}$ . That is,  $w \prec p_{*,\varepsilon}$  for any  $w \in B_\varepsilon(p)$ . In this section, we provide a general construction of  $p_{*,\varepsilon}$ , and show that this property holds. As in the example, the construction will proceed

by forming  $p_{*,\varepsilon}$  by decreasing the smallest entries of  $p$  and increasing the largest entry. Intuitively, one “spreads out” the entries of  $p$  to form  $p_{*,\varepsilon}$  so that  $M_{p_{*,\varepsilon}}$  covers the most area, in order to cover  $B_\varepsilon(p)$ .

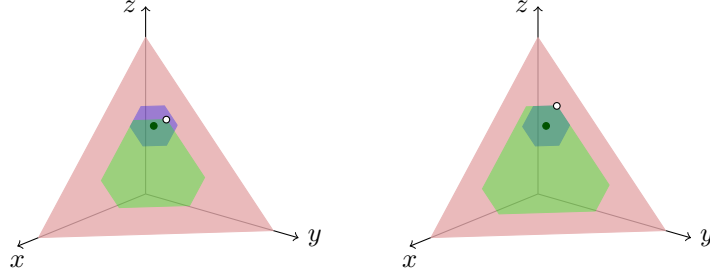


Fig. 3.11 Left: The set  $M_w$  for the point  $w \in B_\varepsilon(p)$  shown in white. Right: The set  $M_{p_{*,\varepsilon}}$  for  $p_{*,\varepsilon}$  the minimizer of  $H$  over  $B_\varepsilon(p)$ .

Let  $\varepsilon > 0$ . We construct a probability vector  $p_{*,\varepsilon}$  which we show has  $M_{p_{*,\varepsilon}} \supseteq B_\varepsilon(p)$  by using the definition of majorization given in (2.3).

**Definition 3.6.1** ( $p_{*,\varepsilon}$ ). If  $p_d^\downarrow > \varepsilon$ , let  $p_{*,\varepsilon}^\downarrow := (p_1^\downarrow + \varepsilon, p_2^\downarrow, \dots, p_d^\downarrow - \varepsilon)$ . Otherwise, let  $\ell \in \{1, \dots, d-1\}$  be the largest index such that the sum of the  $\ell$  smallest entries  $Q_\ell := \sum_{j=d-\ell+1}^d p_j^\downarrow$  has  $Q_\ell \leq \varepsilon$ . If  $\ell = d-1$ , set  $p_{*,\varepsilon}^\downarrow := (1, 0, \dots, 0)$ . Otherwise, choose  $p_{*,\varepsilon}^\downarrow := (p_{j*}^\downarrow)_{j=1}^d$  for

$$p_{j*}^\downarrow := \begin{cases} p_1^\downarrow + \varepsilon & j = 1 \\ p_j^\downarrow & 2 \leq j \leq d - \ell - 1 \\ p_{d-\ell+1}^\downarrow - (\varepsilon - Q_\ell) & j = d - \ell \\ 0 & j \geq d - \ell + 1. \end{cases} \quad (3.40)$$

Note that the condition  $p_d^\downarrow > \varepsilon$  is equivalent to every vector  $w \in B_\varepsilon(p)$  having strictly positive entries. This is the case of Figure 3.3, and  $p_{*,\varepsilon}^\downarrow$  reduces to  $(p_1^\downarrow + \varepsilon, p_2^\downarrow, \dots, p_d^\downarrow - \varepsilon)$ .

Then we set  $p_{*,\varepsilon} := \pi(p_{*,\varepsilon}^\downarrow)$  for the permutation  $\pi \in S_d$  such that  $p = \pi(p^\downarrow)$ . Using Definition 3.6.1, we verify that  $p_{*,\varepsilon} \in B_\varepsilon(p)$ , as follows. First, if  $p_d^\downarrow > \varepsilon$ , then

$$\frac{1}{2} \|p_{*,\varepsilon} - p\| = \frac{1}{2} \|p_{*,\varepsilon}^\downarrow - p^\downarrow\| = \frac{1}{2} (|p_1^\downarrow + \varepsilon - p_1^\downarrow| + |p_d^\downarrow - \varepsilon - p_d^\downarrow|) = \varepsilon.$$

If  $\ell = d-1$ , then  $p_{*,\varepsilon}^\downarrow = (1, 0, \dots, 0)$  and

$$\frac{1}{2} \|p_{*,\varepsilon} - p\| = \frac{1}{2} \|p_{*,\varepsilon}^\downarrow - p^\downarrow\| = \sum_{j=1}^d (p_j^\downarrow - p_{j*}^\downarrow)_+ = \sum_{j=2}^d p_j^\downarrow = Q_\ell \leq \varepsilon$$

yielding  $p_{*,\varepsilon} \in B_\varepsilon(p)$ . Otherwise,  $\left(p_{j*}^\downarrow\right)_{j=1}^d$  is defined via (3.40). For  $j \neq d-\ell$ , that  $p_{j*}^\downarrow \geq 0$  is immediate, and by maximality of  $\ell$ , we have  $p_{(d-\ell)*}^\downarrow = Q_{\ell+1} - \varepsilon > 0$ . Additionally,  $\sum_{j=1}^d p_{j*}^\downarrow = \sum_{j=1}^{d-\ell-1} p_j^\downarrow + Q_\ell - \varepsilon + \varepsilon = \sum_{j=1}^d p_j^\downarrow = 1$ . Furthermore,

$$\begin{aligned} \sum_{j=1}^d |p_{j*}^\downarrow - p_j^\downarrow| &= |p_1^\downarrow + \varepsilon - p_1| + \sum_{j=2}^{d-\ell-1} |p_j^\downarrow - p_j^\downarrow| + |p_{d-\ell}^\downarrow - (\varepsilon - Q_\ell) - p_{d-\ell}^\downarrow| + \sum_{j=d-\ell+1}^d |p_j^\downarrow - 0| \\ &= Q_\ell + (\varepsilon - Q_\ell) + \varepsilon = 2\varepsilon, \end{aligned}$$

so  $p_{*,\varepsilon} \in B_\varepsilon(p)$ .

The following lemma shows that  $p_{*,\varepsilon}$  is indeed the maximal distribution in the majorization order (3.8).

**Lemma 3.6.2.** *We have that  $p_{*,\varepsilon} \succ w$  for any  $w \in B_\varepsilon(p)$ .*

*Proof.* As in Section 3.5, without loss of generality we may take  $p = p^\downarrow$  and  $w = w^\downarrow$ . If  $p_{*,\varepsilon} = (1, 0, \dots, 0)$ , then the result is immediate. If  $p_{*,\varepsilon} = (p_1 + \varepsilon, \dots, p_d - \varepsilon)$ , then consider  $\ell = 0$  and  $Q_0 = 0$  in the following. Now,  $p_{*,\varepsilon} = (p_{j*})$  for  $p_{j*}$  defined via (3.40). Our task is to show that for any  $w = (w_j)_{j=1}^d \in \Delta$ ,

$$\sum_{j=1}^k w_j \leq \sum_{j=1}^k p_{j*} \quad (3.41)$$

holds for each  $k \in \{1, \dots, d-1\}$ . Equality in (3.41) obviously holds for  $k = d$  since  $w, p_{*,\varepsilon} \in \Delta$ .

Since  $\sum_{j=1}^d (w_j - p_j)_+ = \sum_{j=1}^d (w_j - p_j)_- \leq \varepsilon$ , in particular  $\sum_{j=1}^k (w_j - p_j) \leq \varepsilon$  and therefore

$$\sum_{j=1}^k w_j \leq \varepsilon + \sum_{j=1}^k p_j.$$

For  $k \leq d - \ell - 1$ , we have  $\sum_{j=1}^k p_{j*} = \varepsilon + \sum_{j=1}^k p_j$ , yielding (3.41) in this case. On the other hand, for  $k \geq d - \ell$  we have  $\sum_{j=1}^k p_{j*} = \sum_{j=1}^d p_{j*} = 1$ . Since  $\sum_{j=1}^k w_j \leq 1$ , this completes the proof.  $\square$

Given  $\varepsilon > 0$  and a quantum state  $\rho \in \mathcal{D}$ , with eigen-decomposition  $\rho = \sum_{i=1}^d p_i |i\rangle\langle i|$  in the sorted eigenbasis for which  $\rho = \text{Eig}^\downarrow(\rho)$ , we define

$$\rho_{*,\varepsilon} = \sum_{i=1}^d p_{i*} |i\rangle\langle i| \quad (3.42)$$



where  $p_{*,\varepsilon}$  is defined via Definition 3.6.1 where  $p = \vec{\lambda}^\downarrow(\rho)$ . Lemma 3.6.2 therefore proves  $\rho_{*,\varepsilon} \succ \omega$  for any  $\omega \in B_\varepsilon(\rho)$ , proving the second majorization relation of Theorem 3.2.1. The state  $\rho_{*,\varepsilon}$  is unique up to unitary equivalence as follows. If another state  $\tilde{\rho} \in B_\varepsilon(\rho)$  had  $\tilde{\rho} \succ \omega$  for all  $\omega \in B_\varepsilon(\rho)$ , then in particular,  $\tilde{\rho} \succ \rho_{*,\varepsilon} \succ \tilde{\rho}$ , which implies that  $\tilde{\rho}$  and  $\rho_{*,\varepsilon}$  are unitarily equivalent.

### 3.7 Other $p$ -norms

We have established that balls in 1-norm admit a majorization-minimizer and majorization-maximizer. Is that true for other norms?

In the case  $d = 2$ , all compact sets, and in particular  $\varepsilon$ -balls induced by norms, admit a majorization-minimizer. To see this, note that for  $p, q \in \mathcal{P}_2$ , i.e. distributions of the form  $(x, 1 - x)$  for  $x \in [0, 1]$ , we have that  $p \prec q$  if and only if  $\max(p) \leq \max(q)$ . On any compact set  $C \subseteq \mathcal{P}_2$ , the continuous function  $p \mapsto \max(p)$  must attain a minimum which is therefore a majorization-minimizer of  $C$ .

What about for  $d > 2$ ? We show that in dimension  $d = 3$ , for  $p$ -norms, i.e.

$$\|r\|_p := \left( \sum_{i=1}^d r_i^p \right)^{1/p}$$

with  $1 < p < \infty$ , the answer is “no”. To establish this, we will simply compute the majorization-infimum over a ball in  $p$ -norm via (3.20) and show that it lies outside the ball itself.

We denote the  $p$ -norm ball as

$$B_\varepsilon^p(r) := \{q \in \mathcal{P} : \|r - q\|_p \leq \varepsilon\}.$$

**Proposition 3.7.1.** *For any  $1 < p < \infty$  there exists (an uncountable collection of)  $r \in \mathcal{P}_3$  and  $\varepsilon > 0$  such that  $B_\varepsilon^p(r)$  does not admit a majorization-minimizer.*

*Remark.* In the case  $p = \infty$ , the ball  $B_\varepsilon^p(r)$  does appear to admit a majorization-minimizer for any  $d > 2$ ,  $r \in \mathcal{P}_d$  and  $\varepsilon > 0$  (ongoing work).

*Proof.* We will work in arbitrary dimensions  $d > 2$  for now, and only specialize to  $d = 3$  near the end of the proof. For  $\varepsilon > 0$  and  $r \in \mathcal{P}$ , and recall a majorization-infimum  $s$  of

$B_\varepsilon^p(r)$  (which is unique up to permutations) is given by  $s_k = y_k - y_{k-1}$  for

$$\begin{aligned} y_k &:= \text{minimize} && \sum_{i=1}^k q_i^\downarrow \\ &\text{subject to} && q \in B_\varepsilon^p(r) \end{aligned} \tag{3.43}$$

with  $k = 1, \dots, d$  and  $y_0 := 0$ . Define

$$\text{gap}(r) := \min\{|r_i - r_j| : i \neq j, i, j = 1, \dots, d\}.$$

Note that  $\text{gap}(r) = 0$  if  $r_i = r_j$  for some  $i \neq j$ . We have that if  $r \in \mathcal{P}^\downarrow$  and  $\varepsilon < \frac{\text{gap}(r)}{2}$ , then  $B_\varepsilon^p(r) \subseteq \mathcal{P}^\downarrow$ . To see this, note that if  $q \in B_\varepsilon^p(r)$ , then

$$\|q - r\|_\infty \leq \|q - r\|_p \leq \varepsilon < \frac{\text{gap}(r)}{2}$$

and hence each entry of  $q$  must not vary by more than  $\frac{\text{gap}(r)}{2}$ , and thus the entries do not cross and instead remain sorted.

We choose  $r \in \mathcal{P}^\downarrow$  and  $\varepsilon < \frac{\text{gap}(r)}{2}$ . Then (3.43) becomes

$$\begin{aligned} y_k &= \text{minimize} && \sum_{i=1}^k q_i \\ &\text{subject to} && q \in B_\varepsilon^p(r) \end{aligned}$$

or equivalently,

$$\begin{aligned} y_k &= \text{minimize} && \sum_{i=1}^k (r_i + \delta_i) \\ &\text{subject to} && \delta \in \mathbb{R}^n \\ &&& \|\delta\|_p \leq \varepsilon \\ &&& \sum_{i=1}^d \delta_i = 0. \end{aligned}$$

Moreover, while majorization-infimum is non-unique,  $s$  defined via  $s_k = y_k - y_{k-1}$  satisfies  $s \in \mathcal{P}^\downarrow$ . Then for any permutation  $\pi \in S_d$ ,

$$\|s - r\|_p = \|s^\downarrow - r^\downarrow\|_p \leq \|\pi(s) - r\|_p$$

by (3.4) since  $\|\cdot\|_p$  is permutation invariant. Thus, if  $s \notin B_\varepsilon^p(r)$ , then no majorization-infimum of  $B_\varepsilon^p(r)$  lies in the ball. It remains to check that  $\|s - r\|_p > \varepsilon$ .

Next, by substituting  $\delta \rightarrow \varepsilon\delta$ , we obtain for  $k = 1, \dots, d$

$$y_k = \sum_{i=1}^k r_i + \varepsilon\gamma_{k,p,d}$$

where

$$\begin{aligned} \gamma_{k,p,d} &:= \text{minimize} && \sum_{i=1}^k \delta_i \\ &\text{subject to} && \delta \in \mathbb{R}^d \\ &&& \|\delta\|_p \leq 1 \\ &&& \sum_{i=1}^d \delta_i = 0 \end{aligned} \tag{3.44}$$

for  $1 \leq k < d$ , and  $\gamma_{d,p,d} := 0$ . Note that the optimization problem only depends on  $p$ ,  $k$ , and  $d$ , not  $\varepsilon$  or  $r$ . Then the majorization-infimum over  $B_\varepsilon^p(r)$  is given by  $s \in \mathcal{P}^\downarrow$  with

$$s_k := y_k - y_{k-1} = r_k + \varepsilon(\gamma_{k,p,d} - \gamma_{k-1,p,d})$$

using the definitions  $y_0 := 0$  and  $\gamma_{0,p,d} = 0$ . Then

$$\frac{1}{\varepsilon^p} \|r - s\|_p^p = |\gamma_{1,p,d}|^p + \sum_{k=2}^{d-1} |\gamma_{k,p,d} - \gamma_{k-1,p,d}|^p + |\gamma_{d-1,p,d}|^p. \tag{3.45}$$

The right-hand side only depends on  $p$  and  $d$ ; if it can be shown to exceed 1, then it demonstrates the existence of a large family of counterexamples (i.e. corresponding to any  $r \in \mathcal{P}_d^\downarrow$  and  $\varepsilon < \frac{\text{gap}(r)}{2}$ ).

We will compute the dual problem to (3.44). The following is standard; see e.g. [BV04, Chapter 5]. The Lagrangian is given by

$$L(\delta, \lambda, \nu) = \sum_{i=1}^k \delta_i + \lambda(\|\delta\|_p - 1) + \nu \sum_{i=1}^d \delta_i.$$

for  $\delta \in \mathbb{R}^d$ , with dual variables  $\lambda \geq 0$  and  $\nu \in \mathbb{R}$ . Let

$$\Gamma := (\underbrace{\nu + 1, \dots, \nu + 1}_{k \text{ times}}, \nu, \dots, \nu).$$

Then we can write

$$L(\delta, \lambda, \nu) = -\lambda + \langle \Gamma, \delta \rangle + \lambda \|\delta\|_p$$

where  $\langle \Gamma, \delta \rangle := \sum_{i=1}^d \Gamma_i \delta_i$ . Then the dual problem is given by

$$\begin{aligned} & \text{maximize} && g(\lambda, \nu) \\ & \text{subject to} && \lambda \geq 0 \\ & && \nu \in \mathbb{R} \end{aligned}$$

for

$$g(\lambda, \nu) := \min_{\delta \in \mathbb{R}^d} L(\delta, \lambda, \nu).$$

Since there exists a strictly feasible point to the primal problem (e.g.  $\delta = 0_{\mathbb{R}^d}$ ), strong duality holds, and we have  $\gamma_{k,p,d} = \max_{\lambda \geq 0, \nu \in \mathbb{R}} g(\lambda, \nu)$ .

We have that

$$g(\lambda, \nu) = -\lambda + \min_{\delta \in \mathbb{R}^d} [\langle \Gamma, \delta \rangle + \lambda \|\delta\|_p].$$

If  $\lambda = 0$ , then  $g(0, \nu) = -\infty$ . Otherwise, we may change variables  $\delta \rightarrow \lambda \delta$ , yielding

$$g(\lambda, \nu) = -\lambda + \min_{\delta \in \mathbb{R}^d} [\langle \tilde{\Gamma}, \delta \rangle + \|\delta\|_p].$$

for  $\tilde{\Gamma} = \frac{1}{\lambda} \Gamma$ . Then

$$\begin{aligned} g(\lambda, \nu) &= -\lambda - \max_{\delta \in \mathbb{R}^d} [\langle -\tilde{\Gamma}, \delta \rangle - \|\delta\|_p] \\ &= \lambda - f^*(-\tilde{\Gamma}) \end{aligned}$$

where  $f(x) = \|x\|_p$  and  $f^*$  is its conjugate function,

$$f^*(y) := \max_{x \in \mathbb{R}^d} \langle y, x \rangle - f(x).$$

Since  $f$  is a norm, its conjugate is the indicator function of the dual norm (see e.g. [BV04, Example 3.26]),

$$f^*(\tilde{\Gamma}) = \begin{cases} 0 & \|\tilde{\Gamma}\|_q \leq 1 \\ \infty & \text{otherwise} \end{cases}$$

where  $q > 1$  is the Hölder conjugate given by

$$\frac{1}{q} + \frac{1}{p} = 1.$$

Thus,

$$\begin{aligned} g(\lambda, \nu) &= \begin{cases} -\lambda & \lambda > 0 \text{ and } \|\tilde{\Gamma}\|_q \leq 1 \\ -\infty & \text{otherwise,} \end{cases} \\ &= \begin{cases} -\lambda & \lambda > 0 \text{ and } \|\Gamma\|_q \leq \lambda \\ -\infty & \text{otherwise.} \end{cases} \end{aligned}$$

Since  $\|\Gamma\|_q = (k \cdot |\nu + 1|^q + (d - k) \cdot |\nu|^q)^{1/q}$  does not depend on  $\lambda$ , we have that the maximal  $\lambda$  corresponds to  $\lambda = \|\Gamma\|_q$ , and

$$\gamma_{k,p,d} = \max_{\lambda \geq 0, \nu \in \mathbb{R}} g(\lambda, \nu) = -\min_{\nu \in \mathbb{R}} (k \cdot |\nu + 1|^q + (d - k) \cdot |\nu|^q)^{1/q}.$$

The function

$$h(\nu) = k \cdot |\nu + 1|^q + (d - k) \cdot |\nu|^q$$

is differentiable on  $(-\infty, -1) \cup (-1, 0) \cup (0, \infty)$ . Since  $h$  diverges as  $\nu \rightarrow \pm\infty$ , its minimal value occurs either at  $h(0) = k$ ,  $h(-1) = d - k$ , or at some  $\nu \in (-\infty, -1) \cup (-1, 0) \cup (0, \infty)$  with  $h'(\nu) = 0$ . Since

$$h'(\nu) = \begin{cases} -kq|\nu + 1|^{q-1} - (d - k)q|\nu|^{q-1} & \nu < -1 \\ kq|\nu + 1|^{q-1} - (d - k)q|\nu|^{q-1} & -1 < \nu < 0 \\ kq|\nu + 1|^{q-1} + (d - k)q|\nu|^{q-1} & \nu > 0 \end{cases}$$

we have  $h'(\nu) \neq 0$  for  $\nu \notin (-1, 0)$ . Moreover, we find

$$h(\nu^*) = 0 \quad \Longleftrightarrow \quad \nu^* = \frac{-\beta}{\beta + 1}$$

where  $\beta := \left(\frac{k}{d-k}\right)^{\frac{1}{q-1}} > 0$ . Hence,

$$h(\nu^*) = k \left(\frac{1}{1+\beta}\right)^q + (d - k) \left(\frac{1}{1+\beta^{-1}}\right)^q.$$

Next, we show  $h(\nu^*) \leq \min(k, d - k)$  to establish that  $h(\nu^*)$  is the minimum of  $h$ . Note  $\frac{1}{1+\beta} + \frac{1}{1+\beta^{-1}} = 1$ . Let  $t = \frac{1}{1+\beta}$ . We have the following sequence of equivalences

$$kt^q + (d - k)(1 - t)^q \leq d - k$$

$$\begin{aligned}
\beta^{q-1}t^q + (1-t)^q &\leq 1 \\
\beta^{q-1} + (t^{-1} - 1)^q &\leq t^{-q} \\
\beta^{q-1} + \beta^q &\leq (1+\beta)^q \\
\beta^q(\beta+1) &\leq \beta(1+\beta)^q \\
q \log \beta + \log(\beta+1) &\leq \log \beta + q \log(1+\beta) \\
(q-1) \log \beta &\leq (q-1) \log(\beta+1) \\
\beta &\leq \beta+1
\end{aligned}$$

using  $q > 1$ , which shows that  $h(\nu^*) \leq d - k$ . Likewise we have the equivalences

$$\begin{aligned}
kt^q + (d-k)(1-t)^q &\leq k \\
t^q + \beta^{1-q}(1-t)^q &\leq 1 \\
1 + \beta^{1-q}(t^{-1} - 1)^q &\leq t^{-q} \\
1 + \beta &\leq (\beta+1)^q
\end{aligned}$$

where the last line holds since  $q > 1$ . Thus, we have

$$\gamma_{k,p,d} = -h(\nu^*)^{1/q} = - \left( k \left( \frac{1}{1+\beta} \right)^q + (d-k) \left( \frac{1}{1+\beta^{-1}} \right)^q \right)^{1/q} \quad (3.46)$$

for  $\beta = \left( \frac{k}{d-k} \right)^{\frac{1}{q-1}} = \left( \frac{k}{d-k} \right)^{p-1}$ , where  $q$  is the Hölder conjugate to  $p$ .

Together, the relation (3.45) and the formula (3.46) can be used to prove whether or not the majorization-minimizer exists for a given pair  $(p, d)$  with  $1 < p < \infty$  and  $d > 2$ . Let us specialize to the case  $d = 3$ . For  $k = 1$ , we have  $\beta = 2^{-\frac{1}{q-1}}$  and for  $k = 2$  we have  $\beta = 2^{\frac{1}{q-1}}$ . Thus,

$$\gamma_{1,p,3} = \gamma_{2,p,3} = - \left[ \left( 1 + 2^{1/(q-1)} \right)^{-q} + 2 \left( 1 + 2^{1/(1-q)} \right)^{-q} \right]^{1/q}$$

Substituting for  $p$  and simplifying, we obtain

$$|\gamma_{1,p,3}|^p = \frac{2^p}{2 + 2^p}.$$

Thus, the right-hand side of (3.45) becomes  $2 \frac{2^p}{2+2^p}$ . Since

$$2 \frac{2^p}{2 + 2^p} > 1 \iff 2^p > 2 \iff p > 1$$

we have that indeed, the right-hand side of (3.45) exceeds 1 for  $p \in (1, \infty)$ . Note that while  $\lim_{p \rightarrow \infty} 2 \frac{2^p}{2+2^p} = 2 > 1$ , the assumption  $q > 1$  was used earlier in the proof, requiring  $p < \infty$ .  $\square$

*Remark.* If we specialize to the case  $p = 2$ , which implies  $q = 2$ , then (3.46) gives

$$\gamma_{k,2,d} = -\sqrt{\frac{k(d-k)}{d}}.$$

Hence,

$$|\gamma_{1,2,d}|^2 + |\gamma_{d-1,2,d}|^2 = 2 \frac{(d-1)}{d} > 1 \iff d > 2.$$

Thus, the right-hand side of (3.45) is strictly larger than 1 for all  $d > 2$ , when  $p = 2$ .





# Chapter 4

## Majorization flow and continuity bounds

In Section 4.1, we show how Theorem 3.2.1 immediately yields *local* continuity bounds for any Schur concave function. Then in Section 4.2 we introduce the so-called majorization-minimizer map and prove some of its fundamental properties. These are used to establish a notion of “majorization flow” in Section 4.3, which in turn is used to obtain uniform continuity bounds in Section 4.4. This chapter concerns generic techniques applicable to large classes of functions. These techniques and others are used in Chapter 5 to provide continuity bounds for specific functions and families of functions.

### 4.1 Local bounds

As discussed in Chapter 3, Theorem 3.2.1 immediately yields local continuity bounds, as formalized in the following proposition.

**Proposition 4.1.1** (Local continuity bounds). *Let  $f$  be Schur concave on  $\mathcal{P}$ . Let  $p \in \mathcal{P}$  and  $\varepsilon \in [0, 1]$ . Then for any  $q \in B_\varepsilon(p)$ ,*

$$|f(q) - f(p)| \leq \max\{f(p_\varepsilon^*) - f(p), f(p) - f(p_{*,\varepsilon})\}. \quad (4.1)$$

*This result can be restated in the language of quantum states as follows. Let  $f$  be Schur concave on  $\mathcal{D}(\mathcal{H})$ . Let  $\rho \in \mathcal{D}$  and  $\varepsilon \in [0, 1]$ . Then for any state  $\omega \in B_\varepsilon(\rho)$ ,*

$$|f(\omega) - f(\rho)| \leq \max\{f(\rho_\varepsilon^*) - f(\rho), f(\rho) - f(\rho_{*,\varepsilon})\}. \quad (4.2)$$

- Remarks.* 1. Note that this inequality indeed gives a local continuity bound: the right-hand side of (4.1) only depends on  $p$  and  $\varepsilon$ . Moreover, the value of the right-hand side may be computed via the definitions of  $p_\varepsilon^*$  and  $p_{*,\varepsilon}$ .
2. This bound can be used in quite general circumstances because it only requires “oracle-access” to  $f$ , in the sense that one does not need to know the inner workings of  $f$ , but rather only needs three evaluations of it (at  $p$ ,  $p_{*,\varepsilon}$ , and  $p_\varepsilon^*$ ), and the promise that  $f$  is Schur concave. For example, if  $f$  describes the result of an optimization problem, it may be relatively easy to evaluate it (one just needs to solve the problem), but it may not be easy to compute e.g. the derivative of  $f$  along the path of majorization flow,  $\Gamma_f$ , which is introduced in Section 4.3 and is used subsequently to obtain uniform continuity bounds. Note also that  $p_\varepsilon^*$  and  $p_{*,\varepsilon}$  can be computed in time  $O(d \log d)$ , where the time is dominated by sorting the entries of  $p$ . In the quantum case, one additionally needs to compute the eigenvalues of  $\rho$ .

## 4.2 The majorization-minimizer map

By regarding the transformation from  $r \in \mathcal{P}_d$  to the minimal state in majorization order in the  $\varepsilon$ -ball around  $r$ , namely  $r_\varepsilon^*$ , as a map, we can obtain additional information about the behavior of the majorization order on the set of probability vectors. Define the map

$$\begin{aligned} \mathcal{M}_\varepsilon : \quad \mathcal{P}_d &\rightarrow \mathcal{P}_d \\ r &\mapsto r_\varepsilon^*. \end{aligned} \tag{4.3}$$

Note that  $\mathcal{M}_\varepsilon$  is a *non-linear* map on the set of states, and therefore does not represent a physical time evolution. Nevertheless, it has useful mathematical properties.

**Theorem 4.2.1** (Properties of  $\mathcal{M}_\varepsilon$  on  $\mathcal{P}_d$ ). *The (non-linear) map  $\mathcal{M}_\varepsilon$  has following properties, for any  $\varepsilon > 0$ . Let  $r \in \mathcal{P}_d$ .*

- a. *Maps probability vectors to probability vectors:  $\mathcal{M}_\varepsilon : \mathcal{P}_d \rightarrow \mathcal{P}_d$ .*
- b. *Semi-group property: if  $\varepsilon_1, \varepsilon_2 > 0$ , then  $\mathcal{M}_{\varepsilon_1+\varepsilon_2}(r) = \mathcal{M}_{\varepsilon_1} \circ \mathcal{M}_{\varepsilon_2}(r)$ .*
- c. *Covariance with permutations: If  $P$  is a permutation matrix, then  $\mathcal{M}_\varepsilon(Pr) = P\mathcal{M}_\varepsilon(r)$ . In addition,  $\mathcal{M}_\varepsilon(r^\downarrow) \in \mathcal{P}_d^\downarrow$ .*
- d. *Minimal in majorization order:  $\mathcal{M}_\varepsilon(r) \in B_\varepsilon(r)$  and for any  $\omega \in B_\varepsilon(r)$ , we have  $\mathcal{M}_\varepsilon(r) \prec \omega$ .*

- e. *Majorization-preserving*: let  $q \in \mathcal{P}_d$  be such that  $r \prec q$ . Then  $\mathcal{M}_\varepsilon(r) \prec \mathcal{M}_\varepsilon(q)$ .
- f.  $u := (1/d, \dots, 1/d)$  is the unique fixed point of  $\mathcal{M}_\varepsilon$ , i.e. the unique solution to  $r = \mathcal{M}_\varepsilon(r)$  for  $r \in \mathcal{P}_d$ .
- g. For any state  $r \in B_\varepsilon(u)$ , we have  $\mathcal{M}_\varepsilon(r) = u$ . If  $r \notin B_\varepsilon(u)$ , then  $\mathcal{M}_\varepsilon(r)$  is on the boundary of  $B_\varepsilon(r)$ .
- h. For  $p = (1, 0, \dots, 0)$ , the probability vector  $\mathcal{M}_\varepsilon(p)$  is given by

$$\mathcal{M}_\varepsilon(p) = \begin{cases} (1 - \varepsilon, \frac{\varepsilon}{d-1}, \dots, \frac{\varepsilon}{d-1}) & \varepsilon < 1 - \frac{1}{d} \\ u & \varepsilon \geq 1 - \frac{1}{d}. \end{cases} \quad (4.4)$$

- i. *Saturates triangle inequality with  $u$* :

$$\frac{1}{2}\|r - u\|_1 = \frac{1}{2}\|u - \mathcal{M}_\varepsilon(r)\|_1 + \frac{1}{2}\|\mathcal{M}_\varepsilon(r) - r\|_1. \quad (4.5)$$

*Proof.* The properties (a), (c), (g) and (h) follow from the construction given in Section 3.3. Note in particular, that eq. (3.15) establishes the statement that  $r_\varepsilon^*$  lies on the boundary of  $B_\varepsilon(r)$  when  $r_\varepsilon^* \neq u$ . Property (e) can be found in Lemma 2 of [HOS17]. Property (b) is shown by Proposition 4.2.4 below. Property (d) is due to Theorem 3.2.1, and is proven in Section 3.3. The property (f) follows from the fact that the entries of  $\mathcal{M}_\varepsilon(r)$  differ from  $r$  if and only if  $r \neq u$ . Lastly, Property (i) is shown by Proposition 4.2.3 below.  $\square$

Each of these properties has an exact analog in the case of quantum states, which follows from the definition (3.16). For the sake of completeness, these are listed below.

**Theorem 4.2.2** (Properties of  $\mathcal{M}_\varepsilon$  on  $\mathcal{D}(\mathcal{H})$ ). *The (non-linear) map  $\mathcal{M}_\varepsilon$  has following properties, for any  $\varepsilon > 0$ . Let  $\rho \in \mathcal{D}(\mathcal{H})$ .*

- a. *Maps states to states*:  $\mathcal{M}_\varepsilon : \mathcal{D}(\mathcal{H}) \rightarrow \mathcal{D}(\mathcal{H})$ .
- b. *Semi-group property*: if  $\varepsilon_1, \varepsilon_2 > 0$ , we have  $\mathcal{M}_{\varepsilon_1 + \varepsilon_2}(\rho) = \mathcal{M}_{\varepsilon_1} \circ \mathcal{M}_{\varepsilon_2}(\rho)$ .
- c. *Covariance with unitaries*: If  $U$  is a unitary matrix, then  $\mathcal{M}_\varepsilon(U\rho U^*) = U\mathcal{M}_\varepsilon(\rho)U^*$ . In addition,  $\mathcal{M}_\varepsilon(\text{Eig}^\downarrow(\rho))$  is diagonal and sorted.
- d. *Minimal in majorization order*:  $\mathcal{M}_\varepsilon(\rho) \in B_\varepsilon(\rho)$  and for any  $\omega \in B_\varepsilon(\rho)$ , we have  $\mathcal{M}_\varepsilon(\rho) \prec \omega$ .
- e. *Majorization-preserving*: let  $\sigma \in \mathcal{D}(\mathcal{H})$  such that  $\rho \prec \sigma$ . Then  $\mathcal{M}_\varepsilon(\rho) \prec \mathcal{M}_\varepsilon(\sigma)$ .

- f.  $\tau = \frac{1}{d}$  is the unique fixed point of  $\mathcal{M}_\varepsilon$ , i.e. the unique solution to  $\rho = \mathcal{M}_\varepsilon(\rho)$  for  $\rho \in \mathcal{D}(\mathcal{H})$ . Moreover, for any  $\rho \neq \tau$ ,  $\mathcal{M}_\varepsilon(\rho)$  is not unitarily equivalent to  $\rho$ .
- g. For any state  $\rho \in B_\varepsilon(\tau)$ , we have  $\mathcal{M}_\varepsilon(\rho) = \tau$ . If  $\rho \notin B_\varepsilon(\tau)$ , then  $\mathcal{M}_\varepsilon(\rho)$  is on the boundary of  $B_\varepsilon(\rho)$ .
- h. For any pure state  $\psi \in \mathcal{D}_{\text{pure}}(\mathcal{H})$ , we may write  $\psi = \text{diag}(1, 0, \dots, 0)$  in an eigenbasis. In the same basis, the state  $\mathcal{M}_\varepsilon(\psi)$  is given by

$$\mathcal{M}_\varepsilon(\psi) = \begin{cases} \text{diag}(1 - \varepsilon, \frac{\varepsilon}{d-1}, \dots, \frac{\varepsilon}{d-1}) & \varepsilon < 1 - \frac{1}{d} \\ \tau := \frac{1}{d} & \varepsilon \geq 1 - \frac{1}{d}. \end{cases} \quad (4.6)$$

- i. Saturates triangle inequality with  $\tau := \frac{1}{d}$ :

$$\frac{1}{2}\|\rho - \tau\|_1 = \frac{1}{2}\|\tau - \mathcal{M}_\varepsilon(\rho)\|_1 + \frac{1}{2}\|\mathcal{M}_\varepsilon(\rho) - \rho\|_1. \quad (4.7)$$

Two of the properties, namely saturating the triangle inequality with  $u$  and the semi-group property, are proven below. The first, the saturation of the triangle inequality, follows directly from the construction of  $r_\varepsilon^*$ , and proves useful in establishing the second.

**Proposition 4.2.3.** *The state  $r_\varepsilon^*$  saturates the triangle inequality for  $u := (1/d, \dots, 1/d)$  and  $r$ , in that*

$$\frac{1}{2}\|r - u\|_1 = \frac{1}{2}\|u - r_\varepsilon^*\|_1 + \frac{1}{2}\|r_\varepsilon^* - r\|_1.$$

*Proof.* Note

$$\frac{1}{2}\|r_\varepsilon^* - r\|_1 = \min\left(\varepsilon, \frac{1}{2}\|r - u\|_1\right),$$

from the construction of  $r_\varepsilon^*$ . Now, if  $r_\varepsilon^* = u$ , we have the result immediately. Otherwise,

$$\begin{aligned} \|u - r_\varepsilon\|_1 &= m_- \left(\frac{1}{d} - \gamma_-\right) + m_+ \left(\gamma_+ - \frac{1}{d}\right) + \sum_{j=m_++1}^{d-m_-} \left|\frac{1}{d} - r_j^\downarrow\right| \\ &= \frac{m_-}{d} - \sum_{j=1}^{m_+} r_j^\downarrow - \varepsilon + \sum_{j=d-m_-+1}^d r_j^\downarrow - \frac{m_+}{d} - \varepsilon + \sum_{j=m_++1}^{d-m_-} \left|\frac{1}{d} - r_j^\downarrow\right| \\ &= \sum_{j=1}^d \left|\frac{1}{d} - r_j^\downarrow\right| - 2\varepsilon \\ &= \|u - r\|_1 - 2\varepsilon = \|u - r\|_1 - \|r_\varepsilon^* - r\|_1. \end{aligned} \quad \square$$

Next, we establish the semi-group property. This plays a crucial role in the rest of the chapter, and consequently underpins the continuity bounds established in Chapter 5. We highlight that the proof depends heavily on the uniqueness results established in Lemma 3.5.2.

**Proposition 4.2.4.** *If  $\varepsilon_1, \varepsilon_2 > 0$  and  $r \in \mathcal{P}_d$ , we have*

$$\mathcal{M}_{\varepsilon_1+\varepsilon_2}(r) = \mathcal{M}_{\varepsilon_1}(\mathcal{M}_{\varepsilon_2}(r)).$$

*Proof.* Without loss of generality, we may consider  $r = r^\downarrow \in \mathcal{P}^\downarrow$ . Then  $\mathcal{M}_{\varepsilon_1+\varepsilon_2}(r), \mathcal{M}_{\varepsilon_1} \circ \mathcal{M}_{\varepsilon_2}(r), \mathcal{M}_{\varepsilon_2}(r) \in \mathcal{P}^\downarrow$  too.

We first treat the cases in which  $\mathcal{M}_{\varepsilon_1+\varepsilon_2}(r) = u$ ; that is, when  $\frac{1}{2}\|r - u\|_1 \leq \varepsilon_1 + \varepsilon_2$ . If  $\frac{1}{2}\|r - u\|_1 \leq \varepsilon_2$ , then  $\mathcal{M}_{\varepsilon_2}(r) = u$  as well; since  $\mathcal{M}_{\varepsilon_1}(u) = u$ , we have  $\mathcal{M}_{\varepsilon_1+\varepsilon_2}(r) = \mathcal{M}_{\varepsilon_1}(\mathcal{M}_{\varepsilon_2}(r))$ . Next, consider the case  $\varepsilon_2 < \frac{1}{2}\|r - u\|_1 \leq \varepsilon_1 + \varepsilon_2$ . To show  $\mathcal{M}_{\varepsilon_1}(\mathcal{M}_{\varepsilon_2}(r)) = u$ , we need  $\frac{1}{2}\|\mathcal{M}_{\varepsilon_2}(r) - u\|_1 \leq \varepsilon_1$ . By Proposition 4.2.3,

$$\frac{1}{2}\|u - \mathcal{M}_{\varepsilon_2}(r)\|_1 = \frac{1}{2}\|r - u\|_1 - \frac{1}{2}\|\mathcal{M}_{\varepsilon_2}(r) - r\|_1 \leq \varepsilon_1 + \varepsilon_2 - \frac{1}{2}\|\mathcal{M}_{\varepsilon_2}(r) - r\|_1$$

but since  $\frac{1}{2}\|\mathcal{M}_{\varepsilon_2}(r) - r\|_1 = \varepsilon_2$ , using that  $\mathcal{M}_{\varepsilon_2}(r) \neq u$ , we have  $\frac{1}{2}\|u - \mathcal{M}_{\varepsilon_2}(r)\|_1 \leq \varepsilon_1$  as required. This completes the proof of the cases for which  $\mathcal{M}_{\varepsilon_1+\varepsilon_2}(r) = u$ .

For the remainder of the proof, we will consider the construction of the states generated by  $\mathcal{M}_\varepsilon$  via their characterization in Lemma 3.5.2. The uniqueness results of that lemma will allow us to show that  $\mathcal{M}_{\varepsilon_1+\varepsilon_2}(r)$  indeed is the same state as  $\mathcal{M}_{\varepsilon_1} \circ \mathcal{M}_{\varepsilon_2}(r)$ .

Assume that  $\frac{1}{2}\|r - u\|_1 > \varepsilon_1 + \varepsilon_2$ . Let  $r = \sum_{i=1}^d r_i |i\rangle$  and write

$$\mathcal{M}_{\varepsilon_2}(r) = \sum_{i \in I_L} \alpha_- |i\rangle + \sum_{i \in I_M} r_i |i\rangle + \sum_{i \in I_H} \alpha_+ |i\rangle$$

for  $I_L = \{d - m_- + 1, \dots, d\}$ ,  $I_M = \{m_+ + 1, \dots, d - m_-\}$ , and  $I_H = \{1, \dots, m_+\}$ , and  $(\alpha_-, m_-)$  and  $(\alpha_+, m_+)$  are determined by  $r$  and  $\varepsilon_2$  via Lemma 3.5.2.

Now, consider

$$\mathcal{M}_{\varepsilon_1}(\mathcal{M}_{\varepsilon_2}(r)) = \sum_{i \in I_{L'}} \beta_- |i\rangle + \sum_{i \in I_{M'}} r_i |i\rangle + \sum_{i \in I_{H'}} \beta_+ |i\rangle \quad (4.8)$$

with  $I_{L'} = \{d - m'_- + 1, \dots, d\}$ ,  $I_{M'} = \{m'_+ + 1, \dots, d - m'_-\}$ , and  $I_{H'} = \{1, \dots, m'_+\}$ , and where  $(\beta_-, m'_-)$  and  $(\beta_+, m'_+)$  are determined by  $\mathcal{M}_{\varepsilon_2}(r)$  and  $\varepsilon_1$  via Lemma 3.5.2.

We aim to compare the expression (4.8) of  $\mathcal{M}_{\varepsilon_1}(\mathcal{M}_{\varepsilon_2}(r))$  to the following:

$$\mathcal{M}_{\varepsilon_1+\varepsilon_2}(r) = \sum_{i \in I_{L''}} \gamma_- |i\rangle + \sum_{i \in I_{M''}} r_i |i\rangle + \sum_{i \in I_{H''}} \gamma_+ |i\rangle$$

with  $I_{L''} = \{d - m''_- + 1, \dots, d\}$ ,  $I_{M''} = \{m''_+ + 1, \dots, d - m''_-\}$ , and  $I_{H''} = \{1, \dots, m''_+\}$ , and where  $(\gamma_-, m''_-)$  and  $(\gamma_+, m''_+)$  are determined by  $(r, \varepsilon_1 + \varepsilon_2)$  via Lemma 3.5.2. That is, we wish to show  $(\beta_-, m'_-) = (\gamma_-, m''_-)$ , and  $(\beta_+, m'_+) = (\gamma_+, m''_+)$ . We only consider the first equality here; the second is very similar.

Let  $(\nu)_{i=1}^d$  be the entries of  $\mathcal{M}_{\varepsilon_2}(r)$ . That is,  $\nu_i = \alpha_-$  for  $i \in I_L$ ,  $\nu_i = r_i$  for  $i \in I_M$ , and  $\nu_i = \alpha_+$  for  $i \in I_H$ . Equation (3.31) for  $\mathcal{M}_{\varepsilon_1}(\mathcal{M}_{\varepsilon_2}(r))$  in Lemma 3.5.2 yields

$$\nu_{m'_+} > \beta_+ = \frac{1}{m'_+} \left( \sum_{i \in I_{H'}} \nu_i - \varepsilon_1 \right) \geq \nu_{m'_++1}. \quad (4.9)$$

Thus,

- $m'_+ \geq m_+$ . Otherwise,  $\beta_+ = \alpha_+ - \frac{\varepsilon_1}{m'_+} < \alpha_+ = \nu_{m'_++1}$ .
- $m'_+ \leq d - m_-$ . Otherwise,  $\beta_+ < \nu_{m'_+} = \alpha_- < \frac{1}{d}$ , contradicting that  $\beta_+ > \frac{1}{d}$  by Lemma 3.5.2.

Hence,

$$\begin{aligned} \beta_+ &= \frac{1}{m'_+} \left( \sum_{i \in I_H} \nu_i + \sum_{m_++1}^{m'_+} \nu_i - \varepsilon_1 \right) = \frac{1}{m'_+} \left( m_+ \alpha_+ + \sum_{m_++1}^{m'_+} r_i - \varepsilon_1 \right) \\ &= \frac{1}{m'_+} \left( \sum_{i \in I_H} r_i - \varepsilon_2 + \sum_{m_++1}^{m'_+} r_i - \varepsilon_1 \right) = \frac{1}{m'_+} \left( \sum_{i \in I_{H'}} r_i - \varepsilon_1 - \varepsilon_2 \right). \end{aligned}$$

That is,  $\beta_+ = \gamma_+(m'_+)$ . It remains to show  $m'_+ = m''_+$ .

- If  $m'_+ = m_+$ , then  $\nu_{m'_+} = \alpha_+ < r_{m'_+}$  by equation (3.28) for  $\mathcal{M}_{\varepsilon_2}(r)$ .
- If  $m'_+ > m_+$ , then  $\nu_{m'_+} = r_{m'_+}$ .

In either case,  $\nu_{m'_+} \leq r_{m'_+}$ . Hence, (4.9) becomes

$$r_{m_+} \geq \nu_{m'_+} > \beta_+ \geq \nu_{m_++1}.$$

Either  $\nu_{m_++1} = r_{m_++1}$ , or  $\nu_{m_++1} = \alpha_- > r_{m_++1}$ ; in either case,  $\beta_+ \geq r_{m_++1}$ . Thus, writing  $\beta_+ = \gamma_+(m'_+)$ , we have

$$r_{m'_+} > \gamma_+(m'_+) \geq r_{m'_++1}.$$

Equation (3.31) in Lemma 3.5.2 defines  $m''_+$  as the unique solution of

$$r_{m''_+} > \gamma_+(m''_+) \geq r_{m''_++1}$$

and therefore  $m''_+ = m'_+$ . □

### 4.3 Majorization flow

In this section, we consider the infinitesimal action of  $\mathcal{M}_\varepsilon$ . This action is very simple to describe because, as shown in Section 4.3.1, the action of  $\mathcal{M}_\varepsilon$  on  $\mathcal{P}$  is piecewise linear for  $\varepsilon$  small enough. Then in Section 4.3.2, we use the semigroup property to derive integral formulas to track the variation of a quantity along the path traced out by  $(\mathcal{M}_\varepsilon(r))_{\varepsilon \geq 0}$ .

#### 4.3.1 The majorization-minimizer for small $\varepsilon$

**Definition 4.3.1** ( $\delta(r)$ ). Let  $r \in \mathcal{P}$ . If  $r = u \equiv (1/d, \dots, 1/d)$ , then set  $\delta(r) := 0$ . Otherwise, let  $\mu_1 > \mu_2 > \dots > \mu_\ell$  denote the distinct ordered entries of  $r$ , let  $k_+$  be the multiplicity of the largest entry of  $r$  and let  $k_-$  the multiplicity of the smallest entry. Define

$$\delta(r) := \begin{cases} \frac{k_+ k_-}{k_+ + k_-} (\mu_1 - \mu_2) & \ell = 2 \\ \min\{k_+ (\mu_1 - \mu_2), k_- (\mu_{\ell-1} - \mu_\ell)\} & \ell > 2. \end{cases} \quad (4.10)$$

Let us briefly comment on the form of  $\delta(r)$  in the case  $\ell = 2$ . The form used in the definition shows that the value of  $\delta(r)$  is proportional to a gap between consecutive values of  $r$ . For  $\ell = 2$ , however, it can also be written as follows. We have that

$$\text{TV}(r, u) = k_+ \left( r_+ - \frac{1}{d} \right) = k_+ \left( r_+ - \frac{r_+ k_+ + r_- k_-}{k_+ + k_-} \right)$$

using  $k_+ + k_- = d$  and  $r_+ k_+ + r_- k_- = 1$ . Hence,

$$\text{TV}(r, u) = \frac{k_+ r_+ (k_+ + k_-) - r_+ k_+^2 - r_- k_+ k_-}{k_+ + k_-} = \delta(r).$$

For any  $\varepsilon \leq \delta(r)$ , the map  $\mathcal{M}_\varepsilon$  only “moves” the largest and smallest entries of  $r$ , as shown by the following result and illustrated through an example in Figure 4.1.

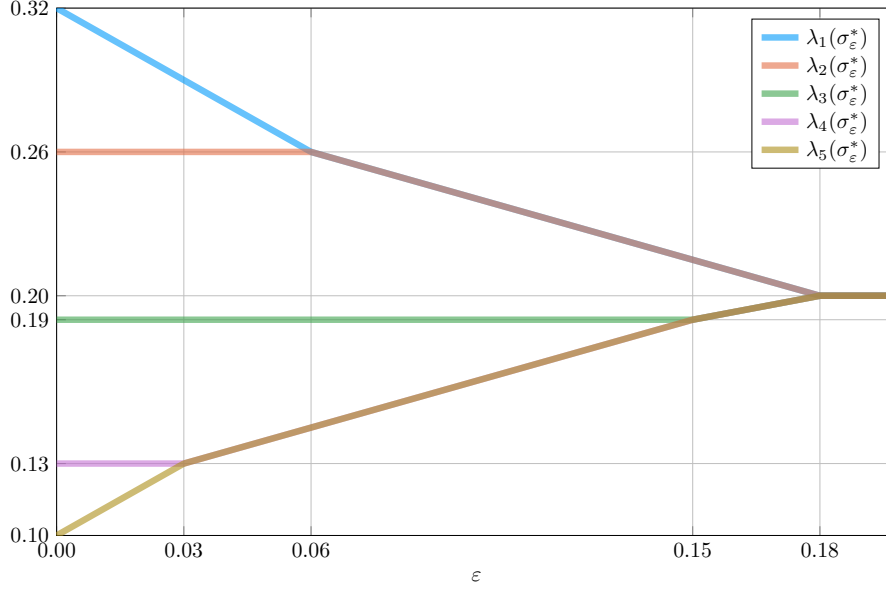


Fig. 4.1 For the 5-dimensional state  $\sigma = \text{diag}(0.32, 0.26, 0.19, 0.13, 0.10)$ , the spectrum of  $\sigma_\varepsilon^* = \mathcal{M}_\varepsilon(\sigma)$  is plotted as a function of  $\varepsilon$ . This plot is a continuous (in  $\varepsilon$ ) analog to the type of plot shown in Figure 3.4, which shows the spectrum of  $\sigma_\varepsilon^*$  at two discrete points,  $\varepsilon = 0$  and  $\varepsilon = 0.07$ , in a different example. Here, at  $\varepsilon = 0$ , the five lines correspond to the five eigenvalues of  $\sigma$ , each with multiplicity one. For  $\varepsilon \leq 0.03$ ,  $\sigma_\varepsilon^* = \text{diag}(0.32 - \varepsilon, 0.26, 0.19, 0.13, 0.10 + \varepsilon)$  and differs from  $\sigma$  only in the smallest and largest eigenvalue. When  $\varepsilon$  reaches 0.03, the multiplicity of the smallest eigenvalue of  $\sigma_\varepsilon^*$  increases to 2. Between  $\varepsilon = 0.03$  and  $\varepsilon = 0.06$ , again only the smallest and largest eigenvalues change, but the smallest eigenvalue has multiplicity 2. This process continues until every eigenvalue reaches  $\frac{1}{d} = 0.2$  at  $T(\sigma, \tau) = 0.18$ .

**Lemma 4.3.2.** *Let  $r \in \mathcal{P}$ . Then  $r_\varepsilon^*$  is a linear perturbation of  $r$  for  $\varepsilon$  small enough, in the sense that*

$$r_\varepsilon^* = r + \varepsilon \mathcal{L}(r) \quad \forall \varepsilon \leq \delta(r) \quad (4.11)$$

where  $\delta(r) \geq 0$  is defined in Definition 4.3.1 and  $\mathcal{L}(r) \in \mathbb{R}^d$  is defined as follows. If  $r = u$  is the uniform distribution, set  $\mathcal{L}(r) := 0$ . For  $r \neq u$ , denote the largest entry of  $r$  as  $r_+$ , and its multiplicity by  $k_+$ , and likewise the smallest entry of  $r$  as  $r_-$  and its multiplicity



by  $k_-$ . Then,  $\mathcal{L}(r) \in \mathbb{R}^n$  is a vector defined by

$$\mathcal{L}(r)_i = \begin{cases} -\frac{1}{k_+} & r_i = r_+ \\ \frac{1}{k_-} & r_i = r_- \\ 0 & \text{else,} \end{cases} \quad (4.12)$$

for  $i = 1, \dots, d$ . Moreover, for any  $r \in \mathcal{P}$ ,

$$\frac{1}{2} \|\mathcal{L}(r)\|_1 = \begin{cases} 0 & r = u, \\ 1 & \text{otherwise,} \end{cases} \quad (4.13)$$

and either

- $r + \delta(r)\mathcal{L}(r) = u$ ,
- $k_+(r + \delta(r)\mathcal{L}(r)) > k_+(r)$ , or
- $k_-(r + \delta(r)\mathcal{L}(r)) > k_-(r)$ .

*Remark.* The last statement shows that the range of validity of Equation (4.11) is exactly until the degeneracy of  $r$  is increased. Since a vector  $r \in \mathcal{P}_d$  has  $d$  entries, the degeneracy can only increase at most  $d - 1$  times, and hence, using the semigroup property, the action of  $\mathcal{M}_\varepsilon$  on  $r \in \mathcal{P}$  can be described as the composition of at most  $d$  linear maps (note that exactly *which* maps depends on  $r$ ;  $\mathcal{M}_\varepsilon$  is not linear on all of  $\mathcal{P}$ ). We used this decomposition in [HD17] to establish uniform continuity bounds for various entropies (the so-called Concave-Type  $(h, \phi)$ -entropies in the language of Section 5.1). However, it has proven useful to consider the action of  $\mathcal{M}_\varepsilon$  fully infinitesimally, dividing it into infinitely many pieces instead of finitely many, as is done in Section 4.3.2. This technique allows us to consider the derivatives of various quantities along the path traced out by  $(\mathcal{M}_\varepsilon(r))_{\varepsilon \geq 0}$  in  $\mathcal{P}$  and subsequently allows us to formulate uniform continuity bounds for a wider class of quantities (e.g. the so-called Convex-Type  $(h, \phi)$ -entropies of Section 5.1, the guesswork without side information considered in Section 6.2, and the expected number of connected components of a particular model of random graph considered in Section 5.5.)

*Proof.* Note that (4.13) follows directly from the definition (4.12). It remains to prove (4.11) and the final statements. We use the notation of Definition 4.3.1. We consider two cases: either  $\varepsilon = \text{TV}(r, u)$ , or  $\varepsilon < \text{TV}(r, u)$ .

Let us consider the case  $\varepsilon = \text{TV}(r, u)$ . First, we establish the inequality  $\delta(r) \leq \text{TV}(r, u)$ . If  $\ell = 2$ , we have that  $\delta(r) = \text{TV}(r, u)$  as discussed after Definition 4.3.1. If  $\ell > 2$ , one cannot have both  $\mu_{\ell-1} > \frac{1}{d}$  and  $\mu_2 < \frac{1}{d}$ . In the case  $\mu_2 \geq \frac{1}{d}$ , we have

$$\delta(r) \leq k_+(\mu_1 - \mu_2) = k_+ \left( \mu_1 - \frac{1}{d} \right) + k_+ \left( \frac{1}{d} - \mu_2 \right) \leq k_+ \left( \mu_1 - \frac{1}{d} \right) \leq \text{TV}(r, u) \quad (4.14)$$

and likewise if  $\mu_{\ell-1} \leq \frac{1}{d}$ ,

$$\delta(r) \leq k_-(\mu_{\ell-1} - \mu_\ell) = k_- \left( \mu_{\ell-1} - \frac{1}{d} \right) + k_- \left( \frac{1}{d} - \mu_\ell \right) \leq k_- \left( \frac{1}{d} - \mu_\ell \right) \leq \text{TV}(r, u). \quad (4.15)$$

Next, let us establish the result in the case  $\varepsilon = \text{TV}(r, u)$ . Note that  $\varepsilon = \text{TV}(r, u)$  necessitates  $\varepsilon = \delta(r) = \text{TV}(r, u)$ . If  $\ell = 2$ , then we directly check that in this case, for  $\varepsilon = \text{TV}(r, u)$ , we have  $r + \varepsilon \mathcal{L} = \mathcal{M}_\varepsilon(r) = u$ , as desired. If  $\ell > 2$  and  $\varepsilon = \text{TV}(r, u)$ , then  $\delta(r) = \text{TV}(r, u)$ . Let us inspect when equality can occur in (4.14) and (4.15). We find that in either case, if  $\delta(r) = \text{TV}(r, u)$  then  $\ell = 3$  and  $\mu_2 = \mu_{\ell-1} = \frac{1}{d}$ , and consequently,  $r + \delta(r) \mathcal{L}(r) = u$ . For example, assume  $\mu_2 \geq \frac{1}{d}$  and equality holds everywhere in (4.15) (the case  $\mu_{\ell-1} \leq \frac{1}{d}$  follows similarly). Then we must have

$$k_+(\mu_1 - \mu_2) \leq k_-(\mu_{\ell-1} - \mu_\ell) \quad (4.16)$$

and  $\mu_2 = \frac{1}{d}$ . Hence, either  $\ell - 1 = 2$  and  $\mu_{\ell-1} = \frac{1}{d}$ , or else  $\ell > 3$  and  $\mu_{\ell-1} < \frac{1}{d}$ . In the latter situation, the right-hand side of (4.16) is strictly less than  $\text{TV}(r, u)$  by following (4.15), while the left-hand side of (4.16) equals  $\text{TV}(r, u)$ , which is a contradiction.

Next, it remains to consider the case  $\varepsilon < \text{TV}(r, u)$ . If  $\ell = 2$ , then since  $\frac{k_+ k_-}{k_+ + k_-} \leq \min(k_+, k_-)$ , we also have

$$\varepsilon \leq \delta(r) \leq \min(k_+(\mu_1 - \mu_2), k_-(\mu_1 - \mu_2)) = \min(k_+(\mu_1 - \mu_2), k_-(\mu_{\ell-1} - \mu_\ell)).$$

The inequality

$$\varepsilon \leq \min(k_+(\mu_1 - \mu_2), k_-(\mu_{\ell-1} - \mu_\ell)). \quad (4.17)$$

therefore holds for any  $\ell \geq 2$  (since it holds by definition of  $\delta(r)$  for  $\ell > 2$ ). Hence, for any  $\ell \geq 2$ , it suffices to prove (4.11) (and that either  $k_+(r + \delta(r) \mathcal{L}(r)) > k_+(r)$  or  $k_-(r + \delta(r) \mathcal{L}(r)) > k_-(r)$ ) under the assumptions that  $\varepsilon < \text{TV}(r, u)$  and (4.17). Here and in the following, we write make the dependence on  $r$  in  $k_\pm$  explicit, writing  $k_\pm \equiv k_\pm(r)$ .

To establish (4.11), we simply check that the choice  $m = k_+(r)$  satisfies the definition of  $m_+(r, \varepsilon)$ , namely that the choice  $m = k_+(r)$  solves (3.10), and likewise for  $k_-$  with  $m_-$ .

To see this, recall that  $\gamma_+^{(m)}(r, \varepsilon) = \frac{1}{m} \left( \sum_{i=1}^m r_i^\downarrow - \varepsilon \right)$ . Then, by taking  $m = k_+(r)$  we find

$$r_{k_+(r)+1}^\downarrow = \mu_2 \leq \frac{1}{k_+(r)} \left( \sum_{i=1}^{k_+(r)} r_i^\downarrow - \varepsilon \right) = \mu_1 - \frac{\varepsilon}{k_+(r)}$$

since  $\frac{\varepsilon}{k_+(r)} \leq \frac{1}{k_+(r)} \delta(r) \leq \mu_1 - \mu_2$ . Additionally,  $\mu_1 - \frac{\varepsilon}{k_+(r)} < \mu_1 = r_{k_+(r)}^\downarrow$ . Therefore,  $m = k_+(r)$  solves (3.10), hence  $m_+(r, \varepsilon) = k_+(r)$ . Proving that  $m_-(r, \varepsilon) = k_-(r)$  is analogous.

Next, consider  $\varepsilon = \delta(r)$ . If  $\ell = 2$ , then  $\varepsilon = \text{TV}(r, u)$  which has already been considered. So consider  $\ell > 3$  and  $\varepsilon < \text{TV}(r, u)$  in the following. Without loss of generality, assume  $\delta(r) = k_+(r)(\mu_1 - \mu_2)$ . We show that  $k_+(\mathcal{M}_\varepsilon(r)) > k_+(r)$ . By the above,  $m_+(r, \varepsilon) = k_+(r)$ , and therefore

$$\gamma_+(r, \varepsilon) = \mu_1 + \frac{\varepsilon}{k_+(r)} = \mu_1 + (\mu_1 - \mu_2) = \mu_2.$$

Hence,  $k_+(r + \delta(r)\mathcal{L}(r)) \geq k_+(r) + k_2 > k_+(r)$  where  $k_2$  is the multiplicity of  $\mu_2$  in  $r$ . Similarly, if  $\delta(r) = k_-(r)(\mu_{\ell-1} - \mu_\ell)$  then  $k_-(\mathcal{M}_\varepsilon(r)) > k_-(r)$ .  $\square$

Let us highlight that the above proof also makes use of the uniqueness results provided by Lemma 3.5.2.

Together, the semigroup property Proposition 4.2.4 and Lemma 4.3.2 provide a simple and effective characterization of  $\mathcal{M}_\varepsilon$  via  $\mathcal{L}$ , which is explored in the next section.

### 4.3.2 The path of majorization flow

For  $r \in \mathcal{P}$ , define the path  $\gamma(t) = \mathcal{M}_t(r)$ . We call the path  $(\gamma(t))_{t=0}^1 \subseteq \mathcal{P}$  as the *path of majorization flow* starting from  $r$ , with respect to the total variation distance. Intuitively speaking, along this path the probability vector decreases in majorization order as quickly as possible, while changing at constant speed in 1-norm. Note that for  $t \geq 1 - \frac{1}{d}$ , one has  $\gamma(t) \equiv u$ , the uniform distribution, and we take the final time as  $t = 1$  for simplicity.

Recall the quantity  $\mathcal{L}$  given in Lemma 4.3.2. Then for  $r \in \mathcal{P}$ ,

$$\mathcal{L}(r) = \partial_s^+ \mathcal{M}_s(r) \Big|_{s=0} \quad (4.18)$$

where  $\partial_s^+$  indicates the one-sided derivative from above. This immediately follows from (4.11). We consequently call  $\mathcal{L}$  the *generator of majorization flow*. The semigroup property established in Proposition 4.2.4 leads to the following integral formulas.

**Theorem 4.3.3.** *Let  $\varepsilon > 0$  and  $(\mathcal{M}_s)_{s \geq 0}$  be as defined above. Then*

$$\mathcal{M}_\varepsilon(r) = r + \int_0^\varepsilon \mathcal{L}(\mathcal{M}_s(r)) \, ds. \quad (4.19)$$

*Additionally, for  $H : \mathcal{P} \rightarrow \mathbb{R}$  which is continuously differentiable on  $\mathcal{P}_+$ , the quantity  $\Gamma_H(r) := \partial_s^+ H(\mathcal{M}_s(r))|_{s=0}$  exists and satisfies*

$$H(\mathcal{M}_\varepsilon(r)) = H(r) + \int_0^\varepsilon \Gamma_H(\mathcal{M}_s(r)) \, ds. \quad (4.20)$$

*Remark.* The relations (4.19) and (4.20) are of particular importance. Equation (4.19) allows one to determine properties of  $\mathcal{M}_s(r)$  by simply analyzing  $\mathcal{L}$ , while (4.20) allows one to analyze continuity properties of  $H$  using  $\Gamma_H$ .

*Proof.* To establish (4.19), note that equation (4.11) immediately yields

$$\partial_\varepsilon^+ \mathcal{M}_\varepsilon(r)|_{\varepsilon=0} = \mathcal{L}(r)$$

where  $\partial_\varepsilon^+$  indicates the one-sided derivative in  $\varepsilon$  from above. Then

$$\partial_s^+ \mathcal{M}_s(r)|_{s=t} = \partial_t^+ \mathcal{M}_{s+t}(r)|_{t=0} = \partial_t^+ \mathcal{M}_t(\mathcal{M}_s(r))|_{t=0} = \mathcal{L}(\mathcal{M}_s(r))$$

follows using the semigroup property. As the path  $s \mapsto \mathcal{M}_s(r)$  is piecewise affine with at most  $d$  pieces, for each fixed  $r \in \mathcal{P}$ , the two-sided derivative  $\partial_s \mathcal{M}_s(r)$  exists for all but at most  $d$  elements  $s \in [0, 1]$ . Then, using the fundamental theorem of calculus,

$$\mathcal{M}_\varepsilon(r) = r + \int_0^\varepsilon \mathcal{L}(\mathcal{M}_s(r)) \, ds.$$

This establishes (4.19). Equation (4.20) follows in the same manner by considering  $H \circ \mathcal{M}_\varepsilon(r)$  instead of  $\mathcal{M}_\varepsilon(r)$ , and using that  $H$  is continuously differentiable on  $\mathcal{P}_+$ .  $\square$

*Remark.* Let us show the power of (4.19) with a quick application. Since  $\frac{1}{2} \|\mathcal{L}(r)\|_1 = 1$  for all  $r \in \mathcal{P} \setminus \{u\}$  as shown in Lemma 4.3.2, and  $\mathcal{L}(u) = 0$ , the triangle inequality

$$\left\| \int_0^\varepsilon \mathcal{L}(\mathcal{M}_s(r)) \, ds \right\|_1 \leq \int_0^\varepsilon \|\mathcal{L}(\mathcal{M}_s(r))\|_1 \, ds$$

immediately yields

$$\frac{1}{2} \|\mathcal{M}_\varepsilon(r) - r\|_1 \leq \varepsilon.$$

This, of course, is not new information;  $\mathcal{M}_\varepsilon(r)$  was constructed in Section 3.3 to have this property. However, we can see that the full construction detailed in that section is no longer needed to derive properties of  $\mathcal{M}_\varepsilon$ , in this case at least. Instead, the simpler construction of  $\mathcal{L}$  (detailed in Lemma 4.3.2) and the semigroup property (as manifested in (4.19)) is enough.

Recall the notation that  $r_\pm$  denotes the largest and smallest elements of  $r \in \mathcal{P}$ .

**Proposition 4.3.4.** *Let  $H : \mathcal{P} \rightarrow \mathbb{R}$  be symmetric<sup>1</sup>, meaning  $H(r) = H(\pi(r))$  for any permutation  $\pi$ , and differentiable on  $\mathcal{P}_+$ . Then for  $r \in \mathcal{P}_+$ ,*

$$\Gamma_H(r) = H_-(r) - H_+(r) \quad (4.21)$$

where  $H_+ = \partial_{r_i} H(r_1, \dots, r_d)$  for any index  $i$  such that  $r_i = r_+$ , and similarly for  $H_-$ . Note, by symmetry, the value of  $H_+$  does not depend on which index  $i$  is chosen, as long as  $r_i = r_+$ .

Moreover, if  $H$  is Schur concave and continuous on  $\mathcal{P}$ , then (4.21) shows that

$$\Gamma_H(r) \geq 0 \quad \forall r \in \mathcal{P} \quad (4.22)$$

using (2.5), and likewise if  $H$  is Schur convex, then  $\Gamma_H(r) \leq 0$  for all  $r \in \mathcal{P}$ .

*Proof.* We have

$$\begin{aligned} \Gamma_H(r) &= \left. \frac{d}{dy} H(\mathcal{M}_y(r)) \right|_{y=0} \\ &= \sum_{i=1}^d \mathcal{L}(r)_i H_i(r) = \sum_{i:r_i=r_+} \frac{-1}{k_+} H_i(r) + \sum_{i:r_i=r_-} \frac{1}{k_-} H_i(r) \end{aligned}$$

where  $H_i(r) = \partial_{r_i} H(r)$ . Since  $H$  is symmetric,  $H_i(r) = H_j(r)$  if  $r_i = r_j$ . Hence, we have (4.21).  $\square$

That is,  $\Gamma_H(r)$  is simply the difference between two partial derivatives of  $H$ , evaluated at  $r$ .

**Example 4.3.5.** Assume  $H : \mathcal{P} \rightarrow \mathbb{R}$  is symmetric. If  $H$  is continuously differentiable on  $\mathcal{P}^\downarrow$  but not on an open set containing  $\mathcal{P}^\downarrow$ , then (4.20) may not hold. For example, define  $H(r) = \sum_{j=1}^d j r_j^\downarrow$ ; this in fact is the *guesswork*, discussed in Chapter 6. Then  $H$  is Schur concave, and for  $r \in \mathcal{P}^\downarrow$ , we have the simple formula  $H(r) = \sum_{j=1}^d j r_j$ . Hence,

---

<sup>1</sup>Note that symmetry is implied by Schur concavity.

$H$  differentiable on  $\mathcal{P}^\downarrow$  with  $\Gamma_H(r) = d - 1$  for  $r \neq u$  and  $\Gamma_H(u) = 0$ . However, setting  $p = (1, 0, \dots, 0)$ , and  $0 < \varepsilon < 1 - \frac{1}{d}$ , we have

$$H(p_\varepsilon^*) = (1 - \varepsilon) + \frac{\varepsilon}{d-1} \sum_{i=2}^d i = 1 + \varepsilon \frac{d}{2},$$

while  $H(p) + \int_0^\varepsilon \Gamma_H(\mathcal{M}_s(p)) ds = 1 + \varepsilon(d-1)$ .

To ameliorate that, we can formulate a slight generalization of Theorem 4.3.3 as follows.

**Corollary 4.3.6.** *Let  $H : \mathcal{P} \rightarrow \mathbb{R}$  be symmetric, and let  $F = H|_{\mathcal{P}^\downarrow}$  have a (possibly non-symmetric) extension  $\tilde{F} : \mathcal{P} \rightarrow \mathbb{R}$  which is differentiable on  $\mathcal{P}_+$ . Then for all  $\varepsilon > 0$ ,*

$$H(\mathcal{M}_\varepsilon(r)) = H(r) + \int_0^\varepsilon \Gamma_{\tilde{F}}(\mathcal{M}_s(r^\downarrow)) ds. \quad (4.23)$$

*Proof.* We have that by Theorem 4.3.3,

$$\tilde{F}(\mathcal{M}_\varepsilon(r)) = \tilde{F}(r) + \int_0^\varepsilon \Gamma_{\tilde{F}}(\mathcal{M}_s(r)) ds. \quad (4.24)$$

Next, since  $H$  is symmetric, we have that

$$H(\mathcal{M}_\varepsilon(r)) = H(\mathcal{M}_\varepsilon(r^\downarrow)) = F(\mathcal{M}_\varepsilon(r^\downarrow)) = \tilde{F}(\mathcal{M}_\varepsilon(r^\downarrow))$$

and similarly  $H(r) = \tilde{F}(r^\downarrow)$ . Hence (4.24) with  $r$  replaced by  $r^\downarrow$  yields the result.  $\square$

Let us see how this repairs Example 4.3.5. We may choose  $\tilde{F}(r) = \sum_{j=1}^d j r_j$  as a differentiable extension of  $H|_{\mathcal{P}^\downarrow}$ . Then  $\Gamma_{\tilde{F}}(r) = \frac{1}{k_-} \sum_{j:r_j=r_-} j - \frac{1}{k_+} \sum_{j:r_j=r_+} j$ , as in the proof of Proposition 4.3.4. Then for  $0 < s < 1 - \frac{1}{d}$  and  $p^\downarrow = p = (1, 0, \dots, 0)$ ,

$$\Gamma_{\tilde{F}}(\mathcal{M}_s(p^\downarrow)) = \Gamma_{\tilde{F}}(1-s, \frac{s}{d-1}, \dots, \frac{s}{d-1}) = \frac{1}{d-1} \sum_{j=2}^d j - 1 = \frac{d}{2}.$$

Hence,

$$H(p) + \int_0^\varepsilon \Gamma_{\tilde{F}}(\mathcal{M}_s(p^\downarrow)) ds = 1 + \varepsilon \frac{d}{2}$$

matching  $H(\mathcal{M}_\varepsilon(p))$ .

**Comparison to other flows** The notion of a flow arises naturally in various branches of physics, mathematics, and engineering. It is interesting to compare and contrast the

notion of majorization flow that we have introduced to the notion of gradient flow that arises in optimal transport and differential geometry, and has been applied to study open quantum systems.

The gradient flow induced by a function  $F$  in a metric space can be loosely interpreted as the flow that decreases  $F$  as quickly as possible [Vil09, p. 645]; in a similar sense, the majorization flow decreases the majorization order as quickly as possible (while following a unit speed path in total variation distance). However, there are several complications that prevent making the connection between majorization flow and gradient flow more precise:

- Perhaps the most obvious one: decreasing in majorization order requires non-increasingness of all the partial sums given in (2.3), while decreasing a function only requires decreasing a scalar value.
- The theory of gradient flow is well-developed on continuous spaces, such as probability measures on  $\mathbb{R}^d$  equipped with a Wasserstein metric (see, e.g. [Vil09]), but is less well-developed in the discrete case considered here (probability measures on  $\{1, \dots, d\}$ ).
- Here we consider the total variation distance, which can be seen as the 1-Wasserstein distance induced by the Hamming distance on the set  $\{1, \dots, d\}$ . Almost all of the literature in discrete or continuous space takes the metric to be the  $p$ -Wasserstein distance for  $p > 1$  (for smoothness reasons).

## 4.4 Uniform continuity bounds from majorization flow

For  $\varepsilon > 0$  and Schur concave  $H$  (meaning  $H(p) \geq H(q)$  if  $q \prec p$ ), we have  $H(\mathcal{M}_\varepsilon(r)) = \max_{p \in B_\varepsilon(r)} H(p)$  since  $\mathcal{M}_\varepsilon(r) = \min_{\prec}(B_\varepsilon(r))$ . In this case, if  $H$  is continuously differentiable on  $\mathcal{P}_+$ , then by using (4.20) we obtain

$$\Delta_\varepsilon^H(r) := \max_{p \in B_\varepsilon(r)} H(p) - H(r) = \int_0^\varepsilon \Gamma_H(\mathcal{M}_s(r)) \, ds. \quad (4.25)$$

In other words, the amount  $H$  can locally increase near  $r$  (quantified by  $\Delta_\varepsilon^H(r)$ ) is determined by  $\Gamma_H$ . Moreover, the global continuity properties of  $H$  are determined by  $\sup_{r \in \mathcal{P}} \Delta_\varepsilon^H(r)$ . To see this, note that if  $p, q \in \mathcal{P}$  satisfy  $\text{TV}(p, q) \leq \varepsilon$ , then

$$H(p) - H(q) \leq H(q_\varepsilon^*) - H(q) = \Delta_\varepsilon^H(q)$$

and

$$H(q) - H(p) \leq H(p_\varepsilon^*) - H(p) = \Delta_\varepsilon^H(p)$$

so we have

$$|H(p) - H(q)| \leq \max\{\Delta_\varepsilon^H(p), \Delta_\varepsilon^H(q)\} \leq \sup_{r \in \mathcal{P}} \Delta_\varepsilon^H(r) \quad (4.26)$$

This fact and (4.25) have two immediate consequences for the continuity properties of  $H$ :

1. If  $\Gamma_H$  is Schur convex, then by (4.25),  $\Delta_\varepsilon^H$  is Schur convex too, as  $\mathcal{M}_s$  is majorization preserving for all  $s \in [0, 1]$ . This provides the upper bound  $\Delta_\varepsilon^H(r) \leq \Delta_\varepsilon^H(\psi)$  for  $\psi = (1, 0, \dots, 0)$ , (since  $\psi$  majorizes every  $r \prec \psi$  for every  $r \in \mathcal{P}$ ) which yields a tight uniform continuity bound by (4.26).
2. If  $\Gamma_H$  can be upper bounded by  $k > 0$  on  $\mathcal{P}_+$ , then by (4.25),  $\Delta_\varepsilon^H(r) \leq \varepsilon k$ . This immediately yields a Lipschitz continuity bound for  $H$  by (4.26).

*Remark.* In case  $H$  is not continuously differentiable on  $\mathcal{P}_+$  but satisfies the conditions of Corollary 4.3.6, the above two points hold with  $\Gamma_H$  replaced by  $\Gamma_{\tilde{F}}$ , in the notation of that corollary.

The second point can be strengthened to the following corollary to Theorem 4.3.3.

**Corollary 4.4.1.** *Let  $H : \mathcal{P} \rightarrow \mathbb{R}$  be a Schur concave function which is continuously differentiable on  $\mathcal{P}_+$ . We write  $H(r_1, \dots, r_d) \equiv H(r)$  for  $r \in \mathcal{P}$ . Next, for  $r \in \mathcal{P}$ , let  $i_+ \in \{1, \dots, d\}$  be an index such that  $r_+ = r_{i_+}$ , and similarly  $i_- \in \{1, \dots, d\}$  such that  $r_- = r_{i_-}$ . Then  $\Gamma_H$  is given by*

$$\begin{aligned} \Gamma_H : \quad \mathcal{P}_+ &\rightarrow \mathbb{R} \\ r &\mapsto (\partial_{r_{i_+}} - \partial_{r_{i_-}})H(r_1, \dots, r_d). \end{aligned} \quad (4.27)$$

*Note that this expression does not depend on the choice of  $i_\pm$  since  $H$  is permutation invariant. Moreover,  $H$  is Lipschitz continuous if and only if*

$$k := \sup_{r \in \mathcal{P}_+} \Gamma_H(r)$$

*satisfies  $k < \infty$ . Moreover, in the latter case  $k$  is the optimal Lipschitz constant for  $H$ .*

*Proof of Corollary 4.4.1.* In the case that  $\sup_{r \in \mathcal{P}_+} \Gamma_H(r) = \infty$ , for each  $n \in \mathbb{N}$ , there exists  $r^{(n)} \in \mathcal{P}_+$  such that

$$\lim_{s \downarrow 0} \frac{H(\mathcal{M}_s(r^{(n)})) - H(r^{(n)})}{s} > n.$$



Additionally,  $r^{(n)} \neq u$ , otherwise the numerator would be zero. Hence, for any  $\varepsilon > 0$  there exists  $0 < s_n < \text{TV}(r^{(n)}, u)$  such that

$$\frac{H(\mathcal{M}_{s_n}(r^{(n)})) - H(r^{(n)})}{s_n} > n - \varepsilon.$$

Since  $s_n < \text{TV}(r^{(n)}, u)$ , we have  $s_n = \text{TV}(\mathcal{M}_{s_n}(r^{(n)}), r^{(n)})$ , and hence

$$\frac{H(\mathcal{M}_{s_n}(r^{(n)})) - H(r^{(n)})}{\text{TV}(\mathcal{M}_{s_n}(r^{(n)}), r^{(n)})} > n - \varepsilon.$$

Any Lipschitz constant  $k$  must be larger than the left-hand side, for any  $n$ , and hence must be  $\infty$ .

Next, consider the case in which  $\sup_{r \in \mathcal{P}_+} \Gamma_H(r)$  is finite. Equation (4.21) shows that the quantity defined in Corollary 4.4.1 is indeed the  $\Gamma_H$  of Theorem 4.3.3. Hence, in the case that  $\sup_{r \in \mathcal{P}_+} \Gamma_H(r) < \infty$  holds, (4.25) and (4.26) show that  $\sup_{r \in \mathcal{P}_+} \Gamma_H(r)$  is indeed a Lipschitz constant for  $H$ . It remains to show this constant is optimal.

Assume there is some Lipschitz constant  $k' < \sup_{r \in \mathcal{P}_+} \Gamma_H(r)$ . For each  $\varepsilon > 0$ , let  $r^{(\varepsilon)} \in \mathcal{P}_+$  satisfy  $\Gamma_H(r^{(\varepsilon)}) > \sup_{r \in \mathcal{P}_+} \Gamma_H(r) - \varepsilon$ . Then

$$k' \geq \frac{H(\mathcal{M}_s(r^{(\varepsilon)})) - H(r^{(\varepsilon)})}{\frac{1}{2} \|\mathcal{M}_s(r^{(\varepsilon)}) - r^{(\varepsilon)}\|_1} \geq \frac{H(\mathcal{M}_s(r^{(\varepsilon)})) - H(r^{(\varepsilon)})}{s}$$

for each  $s > 0$ . Taking the limit  $s \rightarrow 0$  yields

$$k' \geq \Gamma_H(r^{(\varepsilon)}) > \sup_{r \in \mathcal{P}_+} \Gamma_H(r) - \varepsilon.$$

Taking  $\varepsilon \rightarrow 0$  shows  $k' \geq \sup_{r \in \mathcal{P}_+} \Gamma_H(r)$ , which contradicts  $k' < \sup_{r \in \mathcal{P}_+} \Gamma_H(r)$ . Hence, the quantity  $\sup_{r \in \mathcal{P}_+} \Gamma_H(r)$  is indeed the optimal Lipschitz constant.  $\square$



# Chapter 5

## Applications

This section applies the techniques developed in the previous chapters to various tasks (primarily continuity bounds).

- In Section 5.1, tight uniform continuity bounds and optimal Lipschitz constants are obtained for a class of entropies known as  $(h, \phi)$ -entropies.
- Section 5.2 specializes these continuity bounds to the case of the Rényi and Tsallis entropies. In Section 5.2.1, it is shown that the continuity bound for the Shannon entropy (or von Neumann entropy) obtains the best dimensional scaling out of all the  $\alpha$ -Rényi entropies. In Section 5.2.2, majorization flow is used to show why previous uniform continuity bounds for the  $\alpha$ -Rényi entropy with  $\alpha > 1$  achieve poor scaling with  $d$  and  $\alpha$ . Lastly, in Section 5.2.3, the connections between Rényi entropies and thermodynamic free energies introduced by [Bae11] are discussed, and the continuity bounds for the Rényi entropies are interpreted in this framework.
- In Section 5.3, continuity bounds are obtained for several other families of entropies.
- In Section 5.4, continuity bounds are obtained for the number  $K$  of distinct realizations of  $N$  i.i.d. discrete random variables  $X_i \sim p$ , as a function of  $p \in \mathcal{P}_M$  (for some number of possible outcomes  $M$ ).
- In Section 5.5, a continuity bound is obtained for the expected number of connected components of a particular model of random graphs.
- In Section 5.6, the majorization-minimizer and majorization-maximizer are used to provide explicit formulas for smoothed Schur concave and Schur convex functions. The maximum-entropy principle (MaxEnt) is also briefly discussed.

- In Section 5.7, the majorization-minimizer and majorization-maximizer are used to obtain bounds on approximately transforming bipartite pure states via local operations and classical communication.
- In Section 5.8, a simple continuity bound is established for the function  $\rho \mapsto D(\rho\|\sigma)$  on quantum states.

Many of the results established in this chapter are Lipschitz continuity bounds, in which a function  $F$  on  $\mathcal{P}_d$  is shown to satisfy

$$|F(p) - F(q)| \leq k \text{TV}(p, q) \quad \forall p, q \in \mathcal{P}_d$$

for some constant  $k$ , or a function  $F$  on  $\mathcal{D}(\mathcal{H})$  with  $\dim(\mathcal{H}) = d$  is likewise shown to satisfy

$$|F(\rho) - F(\sigma)| \leq k T(\rho, \sigma) \quad \forall \rho, \sigma \in \mathcal{D}(\mathcal{H})$$

for some constant  $k$ . The Lipschitz smallest  $k$  is the optimal Lipschitz constant, and Table 5.1 summarizes the values or bounds on  $k$  established for various quantities in this chapter.

Name	Parameter Regime	Optimal Lipschitz constant $k$	Reference
Tsallis entropy $T_\alpha$	$0 < \alpha < 1$	$k = \infty$	Prop. 5.2.1
	$\alpha > 1$	$k = \frac{\alpha}{\alpha-1}$	Prop. 5.2.1
Rényi entropy $S_\alpha$	$0 < \alpha < 1$	$k = \infty$	Prop. 5.2.2
	$\alpha > 1$	$\frac{\alpha}{\alpha-1} \frac{(d-2)^{1-1/\alpha}}{2 \ln(2)} \leq k \leq \frac{d\alpha}{\alpha-1} \frac{1}{\ln(2)}$	Prop. 5.2.2
	$\alpha = 2$	$k = \begin{cases} \frac{2}{\ln(2)} & d = 2 \\ \frac{d-2}{\sqrt{d-1}-1} \frac{1}{\ln(2)} & d > 2, \end{cases}$	Prop. 5.2.2
	$\alpha = \infty$	$k = \frac{d}{\ln(2)}$	Prop. 5.2.2
von Neumann entropy $S$		$k = \infty$	above (5.26)
$\delta$ -smoothed ( $S^\delta$ )	$0 < \delta < 1$	$k = \log(\delta^{-1} - 1) + \log(d - 1)$	(5.26)
Unified entropies $E_\alpha^s$	$\alpha > 1, s \geq 1$	$k \leq \frac{\alpha}{\alpha-1}$	[HY06]
	$\alpha > 1, s \leq 1$	$k \leq \begin{cases} \frac{\alpha}{\alpha-1} d^{1-\alpha s} & s\alpha < 1 \\ \frac{\alpha}{\alpha-1} & s\alpha \geq 1. \end{cases}$	(5.31)
	$\alpha \in (0, 1)$	$k = \infty$	above (5.31)
$S_f(\rho) = -\text{tr}[f(\rho)]$	$f$ strictly convex $f(0) = f(1) = 0$	$k = f'(1) - f'(0)$	(5.36)
Concurrence $C$		$k = \infty$	(5.37)
Expected number of connected components $E_C$		$k \leq 3 + 0.35\sqrt{d-2}$	(5.42)
Guesswork $G_{\tilde{c}}(X)$		$k = c_d - c_1$	Prop. 6.2.1
with QSI ( $G_{\tilde{c}}(X B)$ )	$c_d < \infty$	$k \leq 2c_d$	Prop. 6.3.3
$\mathbb{E}[K]$	$M < \infty$	$k = N$	(5.38)

Table 5.1 Optimal Lipschitz constants  $k$  for various functions considered in this chapter, as a function of the underlying dimension  $d$  and parameters of the function.  $k = \infty$  means that the function is not Lipschitz continuous.

## 5.1 $(h, \phi)$ -entropies

There are several families of single-partite entropies:  $\alpha$ -Rényi entropies, Tsallis entropies, unified entropies, and so forth (definitions and references for these can be found in Sections 5.2 and 5.3). These various entropies have a lot of similarities, in particular sharing a functional form: for  $r \in \mathcal{P}_d$ ,

$$H(r) = h \left( \sum_{i=1}^d \phi(r_i) \right) \quad (5.1)$$

for a pair of scalar functions  $h$  and  $\phi$ . However, they exhibit different mathematical properties, most notably concavity or lack thereof. For example, the  $\alpha$ -Rényi entropies, defined by

$$S_\alpha(r) = \frac{1}{1-\alpha} \log \left( \sum_{i=1}^d r_i^\alpha \right)$$

are concave for  $\alpha \in (0, 1)$ , but are neither concave nor convex for  $\alpha > 1$ . Certain properties of the latter have proven particularly difficult to study, perhaps as a consequence of the lack of concavity.

While Audenaert proved a tight uniform continuity bound for the  $\alpha$ -Rényi entropies for  $\alpha \in (0, 1)$  in 2007 ([Aud07]; see (5.16) below), a uniform continuity bound on the  $\alpha$ -Rényi entropies for  $\alpha > 1$  was not established until 2011. At that time, Rastegin [Ras11] proved the bound

$$|S_\alpha(p) - S_\alpha(q)| \leq \frac{d^{2(\alpha-1)}}{\alpha-1} [1 - (1-\varepsilon)^\alpha - \varepsilon^\alpha (d-1)^{1-\alpha}] \quad (5.2)$$

where  $\varepsilon = \text{TV}(p, q)$ , for  $\alpha > 1$ . This bound, however, suffers from an exponential dependence on  $\alpha$  (and for fixed  $\alpha$ , polynomial dependence on  $d$ ), while scaling linearly with  $\varepsilon$ , as  $1 - (1-\varepsilon)^\alpha \approx \alpha\varepsilon - \frac{1}{2}\alpha(\alpha-1)\varepsilon^2 + O(\varepsilon^3)$ . Since the inequality  $|S_\alpha(p) - S_\alpha(q)| \leq 2 \log d$  holds trivially, for even moderately large  $\alpha$ , (5.2) provides a non-trivial bound for a very small range of  $\varepsilon$ .

In 2017, Chen et al [CMNF17] improved upon this bound, showing that for  $\alpha > 1$ ,

$$|S_\alpha(p) - S_\alpha(q)| \leq \frac{d^{\alpha-1}}{\alpha-1} [1 - (1-\varepsilon)^\alpha - \varepsilon^\alpha (d-1)^{1-\alpha}] \quad (5.3)$$

However, this bound still suffers from exponential dependence on  $\alpha$ . The proof of both bounds proceeds by reducing to the case of the Tsallis entropy,

$$T_\alpha(p) = \frac{1}{1-\alpha} \left( \left( \sum_{i=1}^d p_i^\alpha \right) - 1 \right),$$

picking up an exponential prefactor along the way.

In this chapter, we prove that the dimensional dependence is at most linear<sup>1</sup>, and in fact

$$|S_\alpha(p) - S_\alpha(q)| \leq \frac{d\alpha}{\alpha - 1} \frac{1}{\ln(2)} \varepsilon \quad (5.4)$$

for  $\text{TV}(p, q) \leq \varepsilon$  and  $\alpha > 1$ . In fact, we prove that the Tsallis entropy increases the fastest near the corners of the probability simplex (i.e. the extremal points which are permutations of  $(1, 0, \dots, 0)$ ) and the slowest near the center of the simplex (see Theorem 5.1.1), while the  $\alpha$ -Rényi entropy increases the fastest close to the center of the probability simplex. This mismatch shows why bounding the difference of Rényi entropies of two probability vectors by the difference of Tsallis entropies of the same two distributions does not work well: a large prefactor is needed to bound the rapidly-changing Rényi entropies near the center of the simplex by the Tsallis entropies which change the slowest there. With the benefit of hindsight (and our proof techniques), we can find that indeed, a linear prefactor suffices to compare the maximum differences in Rényi entropies between two probability vectors which are at a fixed total variation distance apart, and the maximum difference in Tsallis entropies between two distributions at the same distance apart; however, these two maximum differences occur at very different parts of the probability simplex. These two effects can be seen quantitatively in Proposition 5.2.5.

To prove the bound (5.4), as well as determine where each entropy increases the fastest, we take a unified approach to establishing entropic continuity bounds. While concavity only holds for certain entropies, we exploit the fact that all the above entropies  $H$  are Schur concave.

We consider a class of entropic functionals called  $(h, \phi)$ -entropies, which were introduced by [SMMP93]. These are defined by the formula

$$H_{(h, \phi)}(r) := h \left( \sum_{i=1}^d \phi(r_i) \right)$$

---

<sup>1</sup>This is also established in-effect by [WH19, Theorem 7, (2)], which was developed independently and posted slightly later than [HD19] in which we established the linear bound discussed here.

for  $r \in \mathcal{P}_d$ , and likewise by  $H_{(h,\phi)}(\rho) = h(\text{tr}(\phi(\rho)))$  for  $\rho \in \mathcal{D}(\mathcal{H})$ , using the functional calculus (see (10.10)). The concept of  $(h, \phi)$ -entropies for quantum states was introduced by [BZHP+16]. In other words, the  $(h, \phi)$ -entropy of a quantum state is defined by the  $(h, \phi)$ -entropy of the probability vector given by its eigenvalues.

We will consider two classes of  $(h, \phi)$ -entropies which capture almost all single-partite entropies considered in the literature.

- We say a function  $H : \mathcal{P} \rightarrow \mathbb{R}$  is a *Concave-Type  $(h, \phi)$ -entropy* if  $H = H_{(h,\phi)}$  for some  $\phi : [0, 1] \rightarrow \mathbb{R}$  which is continuously differentiable on  $(0, 1]$  and continuous on  $[0, 1]$  and  $h : [\phi(1), \phi(\frac{1}{d})d] \subseteq \mathbb{R} \rightarrow \mathbb{R}$  which is continuously differentiable on  $(\phi(1), \phi(\frac{1}{d})d]$  and continuous on  $[\phi(1), \phi(\frac{1}{d})d]$ , such that  $\phi(0) = 0$  and  $h(\phi(1)) = 0$ , with  $h$  strictly increasing and (not necessarily strictly) concave, and  $\phi$  strictly concave.
- We say a function  $H : \mathcal{P} \rightarrow \mathbb{R}$  is a *Convex-Type  $(h, \phi)$ -entropy* if  $H = H_{(h,\phi)}$  for some  $\phi : [0, 1] \rightarrow \mathbb{R}$  which is continuously differentiable on  $(0, 1]$  and continuous on  $[0, 1]$  and  $h : [\phi(\frac{1}{d})d, \phi(1)] \subseteq \mathbb{R} \rightarrow \mathbb{R}$  which is continuously differentiable on  $(\phi(\frac{1}{d})d, \phi(1))$  and continuous on  $[\phi(\frac{1}{d})d, \phi(1)]$ , such that  $\phi(0) = 0$  and  $h(\phi(1)) = 0$ , with  $h$  strictly decreasing and (not necessarily strictly) convex, and  $\phi$  strictly convex.

Both classes of  $(h, \phi)$ -entropies are strictly Schur concave; this follows immediately from the fact that symmetric and strictly convex (resp. strictly concave) functions are strictly Schur convex (resp. strictly Schur concave), and that the composition with a strictly increasing function preserves strict Schur convexity and strict Schur concavity. In contrast, composition with a strictly decreasing function swaps strict Schur convexity and strict Schur concavity. Likewise, both classes of  $(h, \phi)$ -entropies are continuously differentiable on  $\mathcal{P}_+$  and continuous on  $\mathcal{P}$ .

### Examples of $(h, \phi)$ -entropies

- The  $\alpha$ -Rényi entropy for  $\alpha < 1$ , the  $\alpha$ -Tsallis entropy for  $\alpha > 0$ , the von Neumann entropy (or Shannon entropy in the classical case), the  $(s, \alpha)$ -unified entropies for  $\alpha \in (0, 1)$  and  $s \leq 1$ , and entropies induced by  $f$  divergences with strictly convex  $f$  are all Concave-Type  $(h, \phi)$ -entropies. See also Section 5.4 for an example of a different flavor: there it is shown that  $p \mapsto \mathbb{E}_p[K] - 1$  constitutes a Concave-Type  $(h, \phi)$ -entropy, where  $K$  denotes the number of distinct realizations of  $N$  i.i.d. random variables  $X_i \sim p$  with outcomes in  $\{1, \dots, M\}$ .



- The  $\alpha$ -Rényi entropy for  $\alpha > 1$  and the  $(s, \alpha)$ -unified entropies for  $\alpha > 1$  and  $s \leq 1$  are Convex-Type  $(h, \phi)$ -entropies.

The Rényi and Tsallis entropies are discussed in more detail in Section 5.2, and the other entropies are discussed in Section 5.3.

While both Concave-Type and Convex-Type  $(h, \phi)$ -entropies are Schur concave, Concave-Type  $(h, \phi)$ -entropies are additionally concave, as the composition of a concave increasing function with a concave function. On the other hand, in general Convex-Type  $(h, \phi)$ -entropies are neither convex nor concave. In this chapter, we investigate the continuity properties of these two classes of entropies.

Let  $H = H_{(h, \phi)}$  be an  $(h, \phi)$ -entropy (of either type). By (4.21), we have immediately that

$$\Gamma_H(r) = \left. \frac{d}{dt} H_{(h, \phi)}(\mathcal{M}_t(r)) \right|_{t=0} = h'(\sum_i \phi(r_i))(\phi'(r_-) - \phi'(r_+)) \quad (5.5)$$

where  $r_+$  denotes the largest element of  $r$ , and  $r_-$  the smallest. Note that  $r \mapsto r_-$  is Schur concave, while  $r \mapsto r_+$  is Schur convex.

**Concave-Type  $(h, \phi)$ -entropies** We see that for a Concave-Type  $(h, \phi)$ -entropy  $H \equiv H_{(h, \phi)}$ ,

$$r \mapsto (\phi'(r_-) - \phi'(r_+))$$

is Schur convex and strictly positive for  $r$  non-uniform, and likewise

$$r \mapsto h'(\sum_i \phi(r_i))$$

is strictly positive and strictly Schur convex. Thus,  $\Gamma_H$  is strictly Schur convex on  $\mathcal{P}_+$ , and following the discussion of Section 4.4, we immediately obtain the following results.

**Theorem 5.1.1.** *Let  $H_{(h, \phi)}$  be a Concave-Type  $(h, \phi)$ -entropy. Then  $\Gamma_H$  is strictly Schur convex on  $\mathcal{P}_+$ , and*

$$r \mapsto \sup_{p \in B_\varepsilon(r)} H_{(h, \phi)}(p) - H_{(h, \phi)}(r)$$

*is strictly Schur convex on  $\mathcal{P}$ .*

**Corollary 5.1.2** (Tight uniform continuity bounds for Concave-Type  $(h, \phi)$ -entropies). *For  $\varepsilon > 0$  and any states  $p, q \in \mathcal{P}$  such that  $\text{TV}(p, q) \leq \varepsilon$ , we have*

$$|H_{(h, \phi)}(p) - H_{(h, \phi)}(q)| \leq g(\varepsilon) \quad (5.6)$$

where

$$g(\varepsilon) := \begin{cases} h(\phi(1 - \varepsilon) + (d - 1)\phi(\frac{\varepsilon}{d-1})) & \varepsilon < 1 - \frac{1}{d} \\ h(d\phi(\frac{1}{d})) & \varepsilon \geq 1 - \frac{1}{d} \end{cases} \quad (5.7)$$

and  $d$  is the dimension. Moreover, equality in (5.6) occurs if and only if one of the two distributions (say,  $q$ ) is extremal (i.e. a permutation of  $(1, 0, \dots, 0)$ ), and either

1.  $\varepsilon < 1 - \frac{1}{d}$  and  $p = \pi(1 - \varepsilon, \frac{\varepsilon}{d-1}, \dots, \dots, \frac{\varepsilon}{d-1})$  for some permutation  $\pi$ , or
2.  $\varepsilon \geq 1 - \frac{1}{d}$ , and  $p = u$  is uniform.

This provides a tight uniform continuity bound for the Tsallis entropies, the  $\alpha$ -Rényi entropies for  $\alpha \in (0, 1)$ , the Shannon entropy, the  $(s, \alpha)$ -unified entropies with  $\alpha \in (0, 1)$  and  $s \leq 1$ , and any entropy induced by an  $f$ -divergence or maximal  $f$ -divergence with strictly convex  $f$ . See Section 5.3 for more details and references.

Given an  $(h, \phi)$ -entropy, we may also consider its smoothed variant,

$$H_{(h,\phi)}^\delta(p) := \max_{q \in B_\delta(p)} H_{(h,\phi)}(q) = H_{(h,\phi)} \circ \mathcal{M}_\delta(p) \quad (5.8)$$

for  $\delta \in [0, 1]$ . If  $H_{(h,\phi)}$  is Concave-Type, we can simply establish Lipschitz continuity bounds for any  $\delta > 0$  by using the Schur concavity of  $\Gamma_{H_{(h,\phi)}}$  and Corollary 4.4.1.

**Corollary 5.1.3.** *Let  $H_{(h,\phi)}^\delta$  be the smoothed variant of a Concave-Type  $(h, \phi)$ -entropy, for  $\delta \in [0, 1]$  (as defined in (5.8)). Then  $H_{(h,\phi)}^\delta$  is Lipschitz continuous on  $\mathcal{P}$  if and only if*

$$k := \lim_{\varepsilon \rightarrow 0} \frac{g(\varepsilon + \delta) - g(\delta)}{\varepsilon} < \infty \quad (5.9)$$

where  $g(\varepsilon) := H_{(h,\phi)}(\psi_\varepsilon^*)$  for  $\psi_\varepsilon^* = \text{diag}(1 - \varepsilon, \frac{\varepsilon}{d-1}, \dots, \frac{\varepsilon}{d-1})$  is given in (5.7). Moreover, if  $k$  is finite, then it is the optimal Lipschitz constant for  $H_{(h,\phi)}^\delta$ . In particular, if  $\delta > 0$ , then

$$k = g'(\delta) = h'(\phi(1 - \delta) + (d - 1)\phi(\frac{\delta}{d-1}))(\phi'(\frac{\delta}{d-1}) - \phi'(1 - \delta))$$

and  $H_{(h,\phi)}^\delta$  is Lipschitz continuous.

*Proof.* If  $\delta = 0$ , the result follows from Corollary 5.1.2. Hence, consider the case  $\delta > 0$ . For  $r \in \mathcal{P}_+$ ,

$$\begin{aligned} \Gamma_{H_{(h,\phi)}^\delta}(r) &= \lim_{s \downarrow 0} \frac{1}{s} (H_{(h,\phi)}^\delta(\mathcal{M}_s(r)) - H_{(h,\phi)}^\delta(r)) \\ &= \lim_{s \downarrow 0} \frac{1}{s} (H_{(h,\phi)}(\mathcal{M}_{s+\delta}(r)) - H_{(h,\phi)}(\mathcal{M}_\delta(r))) \end{aligned}$$

$$\begin{aligned}
&= \lim_{s \downarrow 0} \frac{1}{s} \left( H_{(h, \phi)}(\mathcal{M}_s(\mathcal{M}_\delta(r))) - H_{(h, \phi)}(\mathcal{M}_\delta(r)) \right) \\
&= \Gamma_{H_{(h, \phi)}}(\mathcal{M}_\delta(r)).
\end{aligned}$$

Since  $\Gamma_{H_{(h, \phi)}}$  is strictly Schur convex on  $\mathcal{P}_+$  by Theorem 5.1.1, and for any  $\delta > 0$ , the map  $\mathcal{M}_\delta : \mathcal{P} \rightarrow \mathcal{P}_+$  is majorization-preserving, we have that  $r \mapsto \Gamma_{H_{(h, \phi)}}(\mathcal{M}_\delta(r))$  is strictly Schur convex on  $\mathcal{P}$ . Hence, it is maximized at  $r = (1, 0, \dots, 0)$  (or any permutation thereof). Invoking Corollary 4.4.1 completes the proof.  $\square$

**Convex-Type  $(h, \phi)$ -entropies** For a Convex-Type  $(h, \phi)$ -entropy  $H \equiv H_{(h, \phi)}$ ,

$$r \mapsto (\phi'(r_+) - \phi'(r_-))$$

is Schur convex and strictly positive for non-uniform  $r$ , while

$$r \mapsto -h'(\sum_i \phi(r_i))$$

is strictly Schur concave and strictly positive. Hence,  $\Gamma_H$  is the product of a Schur convex and Schur concave function. The former only depends on the largest and smallest entries of  $r$ , however. For  $d = 2$ , these are all the entries, and for  $x \in (0, \frac{1}{2}]$ ,

$$\Gamma_H(\{x, 1 - x\}) = h'(\phi(x) + \phi(1 - x))(\phi'(x) - \phi'(1 - x)) = \text{bin}'_{(h, \phi)}(x)$$

is the derivative of the binary  $(h, \phi)$ -entropy, where

$$\text{bin}_{(h, \phi)}(x) := H_{(h, \phi)}(\{x, 1 - x\}).$$

For  $d > 2$ , define

$$\bar{r} = (r_1, \underbrace{z, \dots, z}_{d-2 \text{ times}}, r_d)$$

where without loss of generality,  $r_1 \geq r_2 \geq \dots \geq r_d$  are the sorted elements of  $r$  and  $z = \frac{1-r_1-r_d}{d-2}$ . Note that  $\bar{r} = Mr$  for

$$M := \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & \frac{1}{d-2} & \cdots & \frac{1}{d-2} & 0 \\ 0 & \vdots & & \vdots & 0 \\ 0 & \frac{1}{d-2} & \cdots & \frac{1}{d-2} & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

which is a doubly-stochastic matrix, and hence the characterization of majorization provided in (2.4) yields that  $\bar{r} \prec r$ . Since  $\bar{r}$  and  $r$  additionally have the same largest and smallest elements, we obtain  $\Gamma_H(r) \leq \Gamma_H(\bar{r})$ . This leads to the following result.

**Theorem 5.1.4.** *Let  $H_{(h,\phi)}$  be a Convex-Type  $(h, \phi)$ -entropy. If  $d = 2$ , then*

$$\sup_{r \in \mathcal{P}_+} \Gamma_H(r) = \sup_{0 < x \leq \frac{1}{2}} \text{bin}'_{(h,\phi)}(x). \quad (5.10)$$

If  $d > 2$ , then

$$\sup_{r \in \mathcal{P}_+} \Gamma_H(r) = \sup_{\substack{x, y: \\ 0 < x \leq \frac{1}{d} \leq y \\ x \leq z \leq y}} -h'(\phi(y) + (d-2)\phi(z) + \phi(x))(\phi'(y) - \phi'(x)), \quad z := \frac{1-y-x}{d-2} \quad (5.11)$$

In either case, if  $\phi$  is differentiable at zero and  $h$  is differentiable at  $\phi(1)$ , then  $H_{(h,\phi)}$  is Lipschitz continuous, and the optimal Lipschitz constant is given by  $\sup_{r \in \mathcal{P}_+} \Gamma_H(r)$ .

*Proof.* The discussion before Theorem 5.1.4 and (5.5) establishes the expressions for  $\sup_{r \in \mathcal{P}_+} \Gamma_H(r)$ , and the proof concludes by Corollary 4.4.1.  $\square$

*Remark.* While (5.10) (resp. (5.11)) do not provide a closed-form expression for  $\sup_{r \in \mathcal{P}_+} \Gamma_H(r)$ , they reduce the naively  $d$ -dimensional optimization problem to a 1- (resp. 2-) dimensional problem.

### 5.1.1 An aside on the necessity of the $(h, \phi)$ form

Theorem 5.1.1 shows Concave-Type  $(h, \phi)$ -entropies  $H$ , which are symmetric and concave, are such that their derivative along the path of majorization flow,  $\Gamma_H$ , is Schur convex. A natural question is whether or not the restriction to the form of an  $(h, \phi)$ -entropy is necessary; perhaps all symmetric and concave functions are such that their derivative along the path of majorization flow is Schur convex? This turns out to not be the case, as shown by Proposition 5.1.6 below. First, we establish a simple characterization of the Schur convexity of  $\Gamma_f$ .

**Lemma 5.1.5.** *Let  $f : \mathcal{P} \rightarrow \mathbb{R}$  be symmetric and concave, and twice differentiable on  $\mathcal{P}_+$ . Then  $\Gamma_f$  is Schur convex on  $\mathcal{P}_+$  iff for each  $x \in \mathcal{P}^\downarrow$ , the Hessian  $H(x)$  of  $f$  at  $x$  satisfies*

$$H(x)_{i,d} + H(x)_{j,1} - H(x)_{j,d} - H(x)_{i,1} \geq 0 \quad (5.12)$$

for each  $i < j$ .

*Proof.* If  $x \in \mathcal{P}^\downarrow$ , then  $\Gamma_f(x) = (\partial_{x_d} - \partial_{x_1})f(x)$ . Equation (2.5) then shows that Schur convexity of  $\Gamma_f$  implies  $(\partial_{x_i} - \partial_{x_j})\Gamma_f(x) \geq 0$  for  $i < j$  and  $x \in \mathcal{P}^\downarrow$ . To recover the reverse implication, one needs to show the assumption (5.12) yields  $(x_i - x_j)(\partial_{x_i} - \partial_{x_j})\Gamma_f(x) \geq 0$  for all  $x \in \mathcal{P}$  and  $i, j$ . We can use that if  $f$  is symmetric, then  $H$  satisfies

$$H(Px) = P^T H(x) P$$

for any permutation matrix  $P$ . Given  $x \in \mathcal{P}$ , let  $P$  be a permutation matrix such that  $x^\downarrow = Px$ . Then applying (5.12) to  $H(x^\downarrow) = P^T H(x) P$  yields the result.  $\square$

Next, we use this characterization to establish a counter-example.

**Proposition 5.1.6.** *There exists a function  $f : \mathcal{P} \rightarrow \mathbb{R}$  which is symmetric, concave, and smooth, but is such that  $\Gamma_f$  is not Schur convex.*

*Proof.* Consider  $d = 4$ ,  $y = (1, 2, 3, 4)$ , and  $z = (4, 3, 2, 1)$ , and define

$$f(x) = -\frac{1}{|S_4|} \sum_{\pi \in S_4} \log(e^{\langle \pi(x), y \rangle} + e^{\langle \pi(x), z \rangle})$$

Note that  $-f(x)$  is the symmetrization of

$$g(x) = \log(e^{\langle x, y \rangle} + e^{\langle x, z \rangle}).$$

Let  $Y$  be a vector-valued random variable which with probability  $1/2$  takes on the value  $2y$  and with probability  $\frac{1}{2}$  takes on the value  $2z$ . Then  $g(x)$  is the cumulant-generating function of  $Y$ , and hence is convex. Thus,  $f(x)$  is concave and symmetric.

Next, we simply check (5.12) with  $i = 2$ ,  $j = 3$ , and  $x = (0.5, 0.3, 0.2, 0) \in \mathcal{P}^\downarrow$ . We find<sup>2</sup> the left-hand side of (5.12) is given by

$$L := -\frac{N(e^{1/5} - 1)^2 e^{1/5}}{D(1 + e^{1/5})^2 (1 + e^{2/5})^2 (1 - e^{1/5} + e^{2/5})^2 (1 + e^{4/5})^2}$$

where

$$\begin{aligned} N := & 3 - 4\sqrt[5]{e} + 13e^{2/5} - 13e^{3/5} + 36e^{4/5} - 32e + 84e^{6/5} - 82e^{7/5} + 157e^{8/5} - 127e^{9/5} \\ & + 235e^2 - 202e^{11/5} + 382e^{12/5} - 290e^{13/5} + 500e^{14/5} - 382e^3 + 685e^{16/5} - 513e^{17/5} \\ & + 817e^{18/5} - 574e^{19/5} + 970e^4 - 680e^{21/5} + 1052e^{22/5} - 691e^{23/5} + 1114e^{24/5} - 756e^5 \end{aligned}$$

---

<sup>2</sup>with some symbolic manipulation by Mathematica; see Appendix A for the code

$$\begin{aligned}
& + 1114e^{26/5} - 691e^{27/5} + 1052e^{28/5} - 680e^{29/5} + 970e^6 - 574e^{31/5} + 817e^{32/5} - 513e^{33/5} \\
& + 685e^{34/5} - 382e^7 + 500e^{36/5} - 290e^{37/5} + 382e^{38/5} - 202e^{39/5} + 235e^8 - 127e^{41/5} \\
& + 157e^{42/5} - 82e^{43/5} + 84e^{44/5} - 32e^9 + 36e^{46/5} - 13e^{47/5} + 13e^{48/5} - 4e^{49/5} + 3e^{10} \\
D := & 3 \left(1 - e^{1/5} + e^{2/5} - e^{3/5} + e^{4/5} - e + e^{6/5}\right)^2 \left(1 + e^{8/5}\right)^2 \left(1 - e^{2/5} + e^{4/5}\right)^2
\end{aligned}$$

which yields  $L \approx -0.15621966760667033050$  and in particular,  $L < 0$ . Hence,  $\Gamma_f$  is not Schur convex.  $\square$

*Remark.* The example used in the proof is for dimension  $d = 4$ . One can likely construct examples in any dimension  $d \geq 4$ , possibly by embedding this example in a larger one. Numerically, I was not able to find a counterexample for  $d = 3$  or lower. One difficulty in constructing such a counterexample is that necessarily two of the indices  $(1, i, j, d)$  must coincide. Perhaps in dimensions  $d = 2$  or  $d = 3$ , symmetric, concave, and differentiable functions  $f$  yield Schur convex  $\Gamma_f$ .

## 5.2 Rényi and Tsallis entropies

Let us apply the results of Corollary 5.1.2 and Theorem 5.1.4 to two cases of particular interest: Rényi and Tsallis entropies. In addition to establishing much sharper continuity bounds for the  $\alpha$ -Rényi entropy for  $\alpha > 1$ , we will also see why previous continuity bounds for the quantity, which were derived from continuity bounds for Tsallis entropies, scale poorly with the dimension.

**Rényi entropies** The  $\alpha$ -Rényi entropies  $S_\alpha$  are parametrized by  $\alpha \in (0, 1) \cup (1, \infty)$ , and are a generalization of the von Neumann entropy in the sense that  $\lim_{\alpha \rightarrow 1} S_\alpha = S$ . The  $\alpha$ -Rényi entropy has been used to bound the quantum communication complexity of distributed information-theoretic tasks [vDH02], can be interpreted in terms of the free energy of a quantum or classical system [Bae11], and is the fundamental quantity defining the entanglement  $\alpha$ -Rényi entropy [WMVF16].

The  $\alpha$ -Rényi entropy [Rén61] for  $\alpha \in (0, 1) \cup (1, \infty)$ , of a quantum state  $\rho \in \mathcal{D}(\mathcal{H})$  is defined by

$$S_\alpha(\rho) := \frac{1}{1 - \alpha} \log \text{tr}(\rho^\alpha).$$

$S_\alpha$  is the  $(h, \phi)$ -entropy with  $h(x) = \frac{1}{1-\alpha} \log x$  for  $x \in \mathbb{R}$  and  $\phi(x) = x^\alpha$  for  $x \in [0, 1]$ . For  $\alpha \in (0, 1)$ ,  $h$  is concave and strictly increasing and  $\phi$  is strictly concave. For  $\alpha > 1$ ,  $h$  is convex and strictly decreasing, and  $\phi$  is strictly convex. Hence,  $S_\alpha$  is a Concave-Type

$(h, \phi)$ -entropy for  $\alpha \in (0, 1)$ , and is a Convex-Type  $(h, \phi)$ -entropy for  $\alpha > 1$ . It is known that  $\lim_{\alpha \rightarrow 1} S_\alpha(\rho) = S(\rho)$ , and  $\lim_{\alpha \rightarrow \infty} S_\alpha(\rho) = S_\infty(\rho) := -\log \lambda_{\max}(\rho)$ , where  $\lambda_{\max}(\rho)$  denotes the largest eigenvalue of  $\rho$ . The Rényi entropies, like the von Neumann (or Shannon) entropy, are additive under tensor products:  $S_\alpha(\rho \otimes \sigma) = S_\alpha(\rho) + S_\alpha(\sigma)$ . In thermodynamic contexts, this property is often called *extensivity*.

For a probability distribution  $p \in \mathcal{P}$ , the above quantity reduces to

$$S_\alpha(p) = \frac{1}{1-\alpha} \log \left( \sum_{i=1}^d p_i^\alpha \right),$$

and  $S_\infty(p) = -\log \max_{1 \leq i \leq d} p_i$ .

**Tsallis entropies** The  $\alpha$ -Tsallis entropy [Tsa88] for  $\alpha \in (0, 1) \cup (1, \infty)$  of a quantum state  $\rho \in \mathcal{D}(\mathcal{H})$  is defined by

$$T_\alpha(\rho) := \frac{1}{1-\alpha} [\text{tr}(\rho^\alpha) - 1],$$

and in the case of a probability vector  $p \in \mathcal{P}$ ,

$$T_\alpha(p) = \frac{1}{1-\alpha} \left[ \sum_{i=1}^d p_i^\alpha - 1 \right].$$

The Tsallis entropy can be seen as a version of the Rényi entropy in which the logarithm has been linearized (up to a factor of  $\ln 2$ ), using the first-order Taylor series  $\log x \approx \frac{1}{\ln 2}(x-1)$ . The Tsallis entropy is not additive under tensor products (it is *nonextensive*) and instead satisfies the relation

$$T_\alpha(\rho \otimes \sigma) = T_\alpha(\rho) + T_\alpha(\sigma) + (1-\alpha)T_\alpha(\rho)T_\alpha(\sigma) \quad (5.13)$$

for  $\alpha \in (0, 1) \cup (1, \infty)$  and  $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ , as can be verified by direct computation.

We have that  $T_\alpha = H_{(h, \phi)}$  for  $h(x) = x$  and  $\phi(x) = \frac{x^\alpha - x}{1-\alpha}$  and hence is a Concave-Type  $(h, \phi)$ -entropy.

**Previously known continuity bounds for Tsallis entropies** Raggio [Rag95, Lemma 2] showed that  $T_\alpha$  is Lipschitz continuous for  $\alpha > 1$ :

$$|T_\alpha(\rho) - T_\alpha(\sigma)| \leq \frac{2\alpha}{\alpha-1} \varepsilon \quad (5.14)$$

if  $T(\rho, \sigma) \leq \varepsilon$ , while Zhang [Zha07b, Theorem 1] proved that if  $T(\rho, \sigma) \leq \varepsilon$  and  $\alpha > 1$ , then

$$|T_\alpha(\rho) - T_\alpha(\sigma)| \leq \begin{cases} \frac{1}{1-\alpha}(\varepsilon^\alpha(d-1)^{1-\alpha} + (1-\varepsilon)^\alpha - 1) & \varepsilon < 1 - \frac{1}{d} \\ \frac{d^{1-\alpha}-1}{(1-\alpha)} & \varepsilon \geq 1 - \frac{1}{d} \end{cases} \quad (5.15)$$

using a coupling technique<sup>3</sup>. In fact, (5.15) also holds in the case  $0 < \alpha < 1$  as was shown by [Aud07, (A.2)] via a direct optimization method (adapting the proof of (1.1)). This bound for all  $\alpha \in (0, 1) \cup (1, \infty)$  also appears as [CMNF17, Lemma 1.2], whose proof appears to follow the same direct optimization method as Audenaert. Zhang [Zha07b, Remark 4] also derived (5.14) from (5.15), and (1.1) from the limit  $\alpha \rightarrow 1$  of (5.15).

[FYK07, Theorem 2.4] showed that if  $\alpha \in [0, 2]$  and  $p, r \in \mathcal{P}$  such that  $\text{TV}(p, r) = \varepsilon \leq \alpha^{1/(1-\alpha)}$ , then

$$|T_\alpha(p) - T_\alpha(r)| \leq (2\varepsilon)^\alpha \ln_\alpha(d) + \eta_\alpha(2\varepsilon)$$

where  $\eta_\alpha(x) = -x^\alpha \ln_\alpha(x)$  and  $\ln_\alpha(x) = \frac{x^{1-\alpha}-1}{1-\alpha}$ . This bound is less tight than (5.15), however.

**Previously known continuity bounds for Rényi entropies** Audenaert [Aud07, Appendix A] proved a tight uniform continuity bound for the  $\alpha$ -Rényi entropies for  $\alpha \in (0, 1)$  in 2007, namely for  $\varepsilon \in [0, 1]$  and  $\rho, \sigma \in \mathcal{D}$  with  $T(\rho, \sigma) \leq \varepsilon$ ,

$$|S_\alpha(\rho) - S_\alpha(\sigma)| \leq \begin{cases} \frac{1}{1-\alpha} \log((1-\varepsilon)^\alpha + (d-1)^{1-\alpha} \varepsilon^\alpha) & \varepsilon < 1 - \frac{1}{d} \\ \log d & \varepsilon \geq 1 - \frac{1}{d}, \end{cases} \quad (5.16)$$

See (5.2) and (5.3) of Section 5.1 for the previously known continuity bounds in the case  $\alpha > 1$ . Note also [WH19, Theorem 7] provides continuity bounds for the  $\alpha$ -Rényi-entropy, although in the case  $\alpha \in (0, 1)$  bounds are not optimal (in contrast to (5.16)), and in the case  $\alpha > 1$ , are not as tight as the bounds presented here.

**New continuity bounds** Corollary 5.1.2 provides an alternate proof of (5.15) for any  $\alpha > 0$  and of (5.16) for  $\alpha \in (0, 1)$  and establishes that in either case for equality to occur, it is necessary and sufficient for one state to be pure, and the other state to have spectrum  $\{1 - \varepsilon, \frac{\varepsilon}{d-1}, \dots, \frac{\varepsilon}{d-1}\}$  if  $\varepsilon < 1 - \frac{1}{d}$ , or  $\{\frac{1}{d}, \dots, \frac{1}{d}\}$  if  $\varepsilon \geq 1 - \frac{1}{d}$ .

---

<sup>3</sup>In fact, [Zha07b, Theorem 1] considers the case  $T(\rho, \sigma) = \varepsilon$ ; in (5.15), their bound has been monotonized to hold for  $T(\rho, \sigma) \leq \varepsilon$ .



**Proposition 5.2.1.** *The  $\alpha$ -Tsallis entropies are Lipschitz continuous for  $\alpha > 1$ , with optimal Lipschitz constant  $\frac{\alpha}{\alpha-1}$ .*

*Proof.*

$$\Gamma_{T_\alpha}(r) = \frac{\alpha}{1-\alpha}(r_-^{\alpha-1} - r_+^{\alpha-1}) = \frac{\alpha}{\alpha-1}(r_+^{\alpha-1} - r_-^{\alpha-1}) \leq \frac{\alpha}{\alpha-1}$$

with equality achieved by  $r = (1, 0, \dots, 0)$ .  $\square$

*Remark.* This improves upon (5.14) by a factor of 2, but can also be derived directly from (5.15).

The following is a corollary of Theorem 5.1.4.

**Proposition 5.2.2.** *The  $\alpha$ -Rényi entropy is Lipschitz continuous if and only if  $\alpha > 1$ . In the latter case, the optimal Lipschitz constant  $k_\alpha$  satisfies*

$$\frac{\alpha}{\alpha-1}(d-2)^{1-1/\alpha} \frac{1}{2 \ln(2)} \leq k_\alpha \leq \frac{d\alpha}{\alpha-1} \frac{1}{\ln(2)}. \quad (5.17)$$

For certain values of  $\alpha$ , we compute  $k_\alpha$  exactly or provide tighter bounds. We have

$$k_\infty = \frac{1}{\ln(2)} d, \quad (5.18)$$

$$k_2 = \begin{cases} \frac{2}{\ln(2)} & d = 2 \\ \frac{d-2}{\sqrt{d-1}-1} \frac{1}{\ln(2)} & d > 2, \end{cases} \quad (5.19)$$

and for  $\alpha \in (1, 2)$ ,

$$k_\alpha \leq \frac{\alpha}{\alpha-1} \frac{d^{\alpha-1}}{\ln(2)} < \frac{d\alpha}{\alpha-1} \frac{1}{\ln(2)}.$$

Additionally, for  $\alpha \in (0, 1)$  and any  $\delta > 0$ , the smoothed entropy  $S_\alpha^\delta$  is Lipschitz continuous, with optimal Lipschitz constant

$$\frac{\alpha}{1-\alpha} \frac{1}{\ln(2)} \frac{(\frac{\delta}{d-1})^{\alpha-1} - (1-\delta)^{\alpha-1}}{(1-\delta)^\alpha + (d-1)^{\alpha-1} \delta^\alpha}.$$

*Remark.* [WH19, Theorem 7 (2)] can be used to establish the bound  $k_\alpha \leq 2 \frac{d\alpha}{\alpha-1} \frac{1}{\ln(2)}$ .

*Proof.* The fact that the  $\alpha$ -Rényi entropies are not Lipschitz for  $\alpha \leq 1$  follows from the fact that  $g(\varepsilon)$  defined in (5.7) has  $\lim_{\varepsilon \rightarrow 0} \frac{g(\varepsilon)}{\varepsilon} = +\infty$ .

Let us prove (5.17); consider  $\alpha > 1$ . First, let's prove the upper bound. We have  $h(x) = \frac{1}{1-\alpha} \log x$  and  $\phi(x) = x^\alpha$ . Then for  $\alpha > 1$ , we have  $\phi(x) \geq 0$  and  $\phi'(x) \geq 0$  for

$x \in [0, 1]$ . Hence, since  $-h'$  is strictly decreasing,

$$\begin{aligned} -h'(\phi(y) + (d-2)\phi(z) + \phi(x))(\phi'(y) - \phi'(x)) &\leq -h'(\phi(y))(\phi'(y) - \phi'(x)) \\ &\leq -h'(\phi(y))\phi'(y) \\ &= -(h \circ \phi)'(y). \end{aligned}$$

Then

$$-(h \circ \phi)'(y) = \frac{\alpha}{\alpha-1} \frac{1}{\ln(2)} \frac{1}{y} \leq \frac{\alpha}{\alpha-1} \frac{1}{\ln(2)} d$$

since  $y \geq \frac{1}{d}$ . Next, consider the lower bound.

Let  $x = 0$ ,  $y = \frac{1}{(d-2)^{1-1/\alpha}} = \frac{1}{(d-2)^{(\alpha-1)/\alpha}}$ ,  $z = \frac{1-y}{d-2}$ , and let  $r = (x, z, \dots, z, y) \in \mathcal{P}^\downarrow$ . Then

$$\begin{aligned} \Gamma_{S_\alpha}(r) &= \frac{\alpha}{\alpha-1} \frac{1}{\ln(2)} \frac{y^{\alpha-1} - x^{\alpha-1}}{x^\alpha + y^\alpha + (d-2)z^\alpha} \\ &= \frac{\alpha}{\alpha-1} \frac{1}{\ln(2)} \frac{y^{\alpha-1}}{y^\alpha + (d-2)z^\alpha} \\ &= \frac{\alpha}{\alpha-1} \frac{1}{\ln(2)} \frac{y^{\alpha-1}}{y^\alpha + (1-y)^\alpha (d-2)^{1-\alpha}} \\ &= \frac{\alpha}{\alpha-1} \frac{1}{\ln(2)} \frac{1}{y} \frac{1}{1 + (1/y - 1)^\alpha (d-2)^{1-\alpha}} \\ &\geq \frac{\alpha}{\alpha-1} \frac{1}{\ln(2)} \frac{1}{y} \frac{1}{1 + y^{-\alpha} (d-2)^{1-\alpha}} \\ &\geq \frac{\alpha}{\alpha-1} \frac{1}{\ln(2)} \frac{1}{y} \frac{1}{1+1} \\ &= \frac{\alpha}{\alpha-1} \frac{1}{2\ln(2)} (d-2)^{1-1/\alpha}. \end{aligned}$$

The proof for  $\alpha = \infty$  is in the next proposition. The proof for  $\alpha \in (1, 2)$  follows from Proposition 5.2.1 and (5.24). Next, for  $\alpha < 1$  and  $\delta > 0$ , we find

$$\lim_{\varepsilon \rightarrow 0} \frac{g(\varepsilon + \delta)}{\varepsilon} = \frac{\alpha}{1-\alpha} \frac{1}{\log_e(2)} \frac{(\frac{\delta}{d-1})^{\alpha-1} - (1-\delta)^{\alpha-1}}{(1-\delta)^\alpha + (d-1)^{\alpha-1} \delta^\alpha}.$$

from which the optimal Lipschitz constant follows by Corollary 5.1.3.

Lastly, we prove establish the optimal Lipschitz constant in the case  $\alpha = 2$ . It suffices to calculate the right-hand side of (5.11) in the case that  $h(x) = -\log x$  and  $\phi(x) = x^2$ . We first establish the result in the case  $d = 2$ , and then in the case  $d > 2$ . For  $d = 2$  and

$\alpha \in (1, 2]$ , we aim to maximize the function

$$h(x, y) = \frac{y^{\alpha-1} - x^{\alpha-1}}{x^\alpha + y^\alpha}$$

where  $y = 1 - x$ . For  $\alpha \in (1, 2]$ , we have the bound

$$\begin{aligned} h(x, y) \leq 1 &\iff y^{\alpha-1} - x^{\alpha-1} \leq y^\alpha + x^\alpha \\ &\iff y^{\alpha-1} - y^\alpha \leq x^{\alpha-1} + x^\alpha \\ &\iff y^{\alpha-1}(1 - y) \leq x^{\alpha-1}(1 - x) \\ &\iff y^{\alpha-1}x \leq x^{\alpha-1}y \\ &\iff y^{\alpha-2} \leq x^{\alpha-2} \end{aligned}$$

using  $y = 1 - x$  and  $x = 1 - y$ . We find  $h(x, y) \leq 1$  since  $x \leq y$  and  $\alpha \in (1, 2]$ . On the other hand,  $h(0, 1) = 1$  for all  $\alpha$ .

For  $d > 2$ , and  $\alpha = 2$ , we consider

$$h(x, y) = \frac{y - x}{x^2 + y^2 + (d - 2)z^2}$$

where  $z \equiv z(x, y) = \frac{1-x-y}{d-2}$ . Then for  $D = x^2 + y^2 + (d - 2)z^2$ ,

$$\begin{aligned} D^2 \frac{d}{dx} h(x, y) &= -(x^2 + y^2 + (d - 2)z^2) - (y - x)(2x - 2z) \\ &= -(x^2 + y^2 + (d - 2)z^2) + 2(y - x)(z - x) \end{aligned}$$

so  $\frac{d}{dx} h(x, y) \leq 0$  iff

$$\begin{aligned} 2(y - x)(z - x) &\leq x^2 + y^2 + (d - 2)z^2 \\ 2(yz - xz + x^2 - xy) &\leq x^2 + y^2 + (d - 2)z^2 \\ 2(yz - xz - xy) &\leq -x^2 + y^2 + (d - 2)z^2 \end{aligned}$$

We have  $2yz \leq y^2 + d^2 \leq y^2 + (d - 2)z^2$ , so it remains to show  $-2(xz + xy) \leq -x^2$ . But that follows from

$$x^2 \leq 4x^2 \leq 2(xz + xy)$$

using  $x \leq y$  and  $x \leq z$ . Thus, for any  $y \geq x$ , such that  $x \leq z \leq y$ ,  $h(x, y)$  is decreasing in  $x$ .

Hence, we consider  $x = 0$ . Then  $z = \frac{1-y}{d-2}$ , and  $y$  is constrained only by  $y \in [\frac{1}{d-1}, 1]$ . Then

$$h(0, y) = \frac{y}{y^2 + (d-2)^{-1}(1-y)^2} = \frac{y(d-2)}{(d-2)y^2 + 1 - 2y + y^2} = \frac{(d-2)y}{(d-1)y^2 + 1 - 2y}.$$

We have

$$((d-1)y^2 + 1 - 2y)^2 \partial_y h(0, y) = ((d-1)y^2 + 1 - 2y)(d-2) - (d-2)y(2(d-1)y - 2)$$

so  $\partial_y h(0, y) = 0$  if and only if

$$\begin{aligned} (d-1)y^2 + 1 - 2y &= y(2(d-1)y - 2) \\ (d-1)y^2 + 1 - 2y &= 2(d-1)y^2 - 2y \\ 1 &= (d-1)y^2 \\ y &= \frac{1}{\sqrt{d-1}}. \end{aligned}$$

Note that with  $x = 0$ ,  $y = \frac{1}{\sqrt{d-1}}$ , we have  $0 \leq x \leq z \leq y \leq 1$  so the constraints are satisfied. It could be that this choice of  $y$  yields only a local maximum. To rule this case out, since  $y \in [\frac{1}{d-1}, 1]$ , we check  $h(0, \frac{1}{d-1}) = h(0, 1) = 1$ , while

$$h(0, (d-1)^{-1/2}) = \frac{d-2}{2\sqrt{d-1}-2} > 1$$

for  $d \geq 3$ . □

For  $\alpha = \infty$ , we can obtain both a tight uniform continuity bound and the optimal Lipschitz continuity constant.

**Proposition 5.2.3.** *For  $\varepsilon > 0$  and  $p, q \in \mathcal{P}_d$  such that  $\text{TV}(p, q) \leq \varepsilon$ , we have  $|S_\infty(p) - S_\infty(q)| \leq \log(1 + \varepsilon d)$  if  $\text{TV}(p, q) \leq \varepsilon$ . In particular,  $S_\infty$  has an optimal Lipschitz constant of  $\frac{d}{\ln(2)}$ .*

*Remark.* [WH19, Lemma 18] provides a simple direct proof to establish the Lipschitz constant  $\frac{d}{\ln(2)}$ .

*Proof.* Since  $S_\infty(r) = -\log r_+$ , for  $r \neq u$ , (4.21) yields

$$\Gamma_{S_\infty}(r) = (\partial_- - \partial_+)S_\infty(r) = \partial_+ \log r_+ = \frac{1}{\ln(2)} \frac{1}{r_+},$$

whereas  $\Gamma_{S_\infty}(u) = 0$  as  $\mathcal{L}(u) = 0$ . The optimal Lipschitz constant follows from the fact that  $r_+ \geq \frac{1}{d}$  and for some probability vectors,  $r_+ = \frac{1}{d}$ .

Moreover,  $r \mapsto \frac{1}{\ln(2)} \frac{1}{r_+}$  is Schur concave (as the composition of a decreasing function and the Schur convex function  $r \mapsto r_+$ ). Hence,  $\Gamma_{S_\infty}$  is Schur concave on  $\mathcal{P} \setminus \{u\}$ . Since  $\mathcal{M}_s$  is majorization-preserving for all  $s \in [0, 1]$ , we have

$$r \mapsto \Delta_\varepsilon^{S_\infty} := S_\infty(\mathcal{M}_\varepsilon(r)) - r = \int_0^\varepsilon \Gamma_{S_\infty}(\mathcal{M}_s(r)) \, ds$$

is Schur concave on  $\mathcal{P} \setminus B_\varepsilon(u)$ . This uses the fact that for  $r \in \mathcal{P} \setminus B_\varepsilon(u)$ ,  $\text{TV}(r, u) > \varepsilon$  and hence  $\mathcal{M}_s(r) \neq u$  for all  $s \in [0, \varepsilon]$ . For any  $r \in \mathcal{P} \setminus B_\varepsilon(u)$ ,  $\mathcal{M}_t(r) \prec r$  for  $t = \text{TV}(r, u)$ . Hence,

$$\max_{r \in \mathcal{P}} S_\infty(\mathcal{M}_\varepsilon(r)) - r = \max_{r \in B_\varepsilon(u)} S_\infty(\mathcal{M}_\varepsilon(r)) - r = \max_{r \in B_\varepsilon(u)} \log(d) - S_\infty(r)$$

using that  $\mathcal{M}_\varepsilon(r) = u$  for  $r \in B_\varepsilon(u)$ . Then

$$\max_{r \in B_\varepsilon(u)} \log(d) - S_\infty(r) = \max_{r \in B_\varepsilon(u)} \log(dr_+).$$

For  $r \in B_\varepsilon(u)$ ,  $r_+ \leq \frac{1}{d} + \varepsilon$ , with equality for  $r = (\frac{1}{d} + \varepsilon, \frac{1}{d}, \dots, \frac{1}{d}, \frac{1}{d} - \varepsilon)$ . Hence, putting it all together,

$$\max_{r \in \mathcal{P}} S_\infty(\mathcal{M}_\varepsilon(r)) - r = \log \left( d \left( \frac{1}{d} + \varepsilon \right) \right) = \log(1 + d\varepsilon). \quad \square$$

*Remark.* The proof of Proposition 5.2.3 shows the Schur concavity of  $\Delta_\varepsilon^{S_\infty}$  on  $\mathcal{P} \setminus B_\varepsilon(u)$ . This contrasts strongly with the Schur *convexity* of  $\Delta_\varepsilon^H$  on  $\mathcal{P}$  for all Concave-Type  $(h, \phi)$  entropies  $H$  proven in Theorem 5.1.1.

### 5.2.1 Rényi entropy of parameter $\alpha = 1$ has optimal dimensional scaling in its continuity bound

Note that the Rényi entropy of parameter  $\alpha = 1$  is the Shannon entropy (or von Neumann entropy, in the quantum case). There is a sense in which the continuity properties as a function of dimension  $d$  of the Rényi entropy  $S_\alpha$  are much improved at  $\alpha = 1$  compared to  $\alpha \neq 1$ . Let us introduce some notation. Define

$$C_\alpha(d, \varepsilon) := \sup_{\substack{p, q \in \mathcal{P}_d \\ \text{TV}(p, q) \leq \varepsilon}} |S_\alpha(p) - S_\alpha(q)|$$

as the optimal uniform continuity bound for  $S_\alpha$  over  $\mathcal{P}_d$ . Consider a sequence  $(\varepsilon_d)_{d \in \mathbb{N}}$  such that  $\varepsilon_d \xrightarrow{d \rightarrow \infty} 0$ . Clearly, if

$$\limsup_{d \rightarrow \infty} C_\alpha(d, \varepsilon_d) = 0 \quad (5.20)$$

then for any sequences of distributions  $(p_d)_{d \in \mathbb{N}}$  and  $(q_d)_{d \in \mathbb{N}}$  with  $p_d, q_d \in \mathcal{P}_d$  such that  $\text{TV}(p_d, q_d) \leq \varepsilon_d$ , we have

$$\limsup_{d \rightarrow \infty} |S_\alpha(p_d) - S_\alpha(q_d)| = 0.$$

Thus, any  $(\varepsilon_d)_{d \in \mathbb{N}}$  satisfying (5.20) provides a dimensionally-aware notion of continuity for  $S_\alpha$ . Moreover, the slower  $\varepsilon_d$  converges to zero, the stronger the statement of continuity provided by (5.20). The following proposition therefore demonstrates that the case  $\alpha = 1$  is the “most continuous” in this sense.

**Proposition 5.2.4.** *For any  $s > 0$ ,*

$$C_1(d, d^{-s}) \xrightarrow{d \rightarrow \infty} 0. \quad (5.21)$$

*In fact, if  $(\varepsilon_d)_{d \in \mathbb{N}}$  has  $\varepsilon_d \log(d) \rightarrow 0$ , then  $\lim_{d \rightarrow \infty} C_1(d, \varepsilon_d) = 0$ . However, for  $\alpha \in (0, 1) \cup (1, \infty)$ ,*

$$\liminf_{d \rightarrow \infty} C_\alpha(d, d^{-\frac{|\alpha-1|}{\alpha}}) > 0. \quad (5.22)$$

*and likewise  $\liminf_{d \rightarrow \infty} C_\infty(d, d^{-1}) > 0$ .*

*Remark.* Note that this result contrasts with the scenario of fixed dimension  $d$ . The fact that  $S_\alpha$  is Lipschitz continuous on  $\mathcal{P}_d$  if and only if  $\alpha > 1$  provides a notion in which  $S_\alpha$  is “more continuous” for  $\alpha > 1$  than for  $\alpha > 0$ . In other words, if  $\alpha > 0$ , the only constant  $k$  satisfying  $|S_\alpha(p) - S_\alpha(q)| \leq k \text{TV}(p, q)$  is  $k = \infty$ , while for  $\alpha > 1$ , finite  $k$  suffices. This notion is not “dimensionally-aware”, however, in the sense that  $d$  is fixed.

Additionally, the parameter  $\alpha = 1$  for the Tsallis entropy, which again coincides with the Shannon entropy, does not admit optimal scaling of the continuity bound out of the whole family of Tsallis entropies; in fact, (5.15) shows that the optimal bound actually decreases with dimension for  $\alpha > 1$ . This is related to the nonextensivity of the Tsallis entropy; extensivity enforces the scaling

$$S_\alpha(p^{\otimes n}) - S_\alpha(q^{\otimes n}) = n(S_\alpha(p) - S_\alpha(q))$$

and since

$$\text{TV}(p^{\otimes n}, q^{\otimes n}) \leq n \text{TV}(p, q)$$

we find that

$$C_\alpha(d^n, n\varepsilon_d) \geq nC_\alpha(d, \varepsilon_d),$$

or, for  $d = 2^n$ ,

$$C_\alpha(d, \varepsilon_d \log(d)) \geq \log(d) C_\alpha(2, \varepsilon_d).$$

Hence for any extensive entropy, the optimal continuity bound must grow at least logarithmically with dimension (up to the modification  $\varepsilon_d \rightsquigarrow \log(d)\varepsilon_d$ ).

*Proof.* For  $0 \leq \varepsilon < 1 - \frac{1}{d}$ , the Audenaert-Fannes bound (1.1) gives

$$C_1(d, \varepsilon) = h_2(\varepsilon) + \varepsilon \log(d-1)$$

where  $h_2(\varepsilon) = -\varepsilon \log(\varepsilon) - (1-\varepsilon) \log(1-\varepsilon)$  is the binary entropy. Hence, if  $\varepsilon_d \log(d) \rightarrow 0$ , then  $\varepsilon_d \rightarrow 0$  and

$$C_1(d, \varepsilon_d) = h_2(\varepsilon_d) + \varepsilon_d \log(d-1) \rightarrow 0$$

as well. Since for any  $s > 0$  we have  $d^{-s} \log(d) \rightarrow 0$ , we recover (5.21).

Next, let us establish (5.22) for  $0 < \alpha < 1$ . In this case, for  $0 \leq \varepsilon < 1 - \frac{1}{d}$  (5.16) gives

$$C_\alpha(d, \varepsilon) = \frac{1}{1-\alpha} \log((1-\varepsilon)^\alpha + (d-1)^{1-\alpha} \varepsilon^\alpha).$$

For  $\varepsilon_d = d^{-\frac{1-\alpha}{\alpha}}$ , we have  $(d-1)^{1-\alpha} \varepsilon_d^\alpha \rightarrow 1$  while  $(1-\varepsilon_d)^\alpha \rightarrow 0$ , and hence  $C_\alpha(d, \varepsilon_d) \rightarrow \frac{1}{1-\alpha} > 0$ .

For  $\alpha > 1$ , we do not have an exact expression for  $C_\alpha(d, \varepsilon)$ . However, since a Lipschitz bound provides a uniform continuity bound, we have

$$C_\alpha(d, \varepsilon) \geq \varepsilon \sup_{r \in \mathcal{P}_+} \Gamma_{S_\alpha}(r)$$

using Corollary 4.4.1. Then Proposition 5.2.2 shows that  $k_\alpha = \sup_{r \in \mathcal{P}_+} \Gamma_{S_\alpha}(r)$  satisfies

$$k_\alpha \geq \frac{\alpha}{\alpha-1} (d-2)^{1-1/\alpha} \frac{1}{2 \ln(2)}.$$

Hence, if  $\varepsilon_d = d^{-\frac{\alpha-1}{\alpha}}$ , then

$$\varepsilon_d (d-2)^{1-1/\alpha} = \varepsilon_d (d-2)^{(1-\alpha)/\alpha} \rightarrow 1.$$

Therefore  $\liminf_{d \rightarrow \infty} C_\alpha(d, \varepsilon) \geq \frac{\alpha}{\alpha-1} \frac{1}{2 \ln(2)} > 0$ . Likewise, for  $\alpha = \infty$ , the optimal Lipschitz constant is given in Proposition 5.2.3 as  $k_\infty = \frac{d}{\ln(2)}$ , and hence  $\liminf_{d \rightarrow \infty} C_\infty(d, d^{-1}) \geq \frac{1}{\ln(2)} > 0$ .  $\square$

### 5.2.2 Discussion of previous continuity bounds for $S_\alpha$ with $\alpha > 1$

As mentioned at the start of the Section 5.1, continuity bounds on the  $\alpha$ -Rényi entropy for  $\alpha > 1$  were not known until 2011, and the bounds known until this work have poor scaling  $\sim d^{\alpha-1}$ . Here, we use majorization flow as a tool to understand why the previous bounds performed poorly. The technique used to establish the previous bounds was to relate the difference in Rényi entropy of two distributions to the corresponding difference in Tsallis entropy [Ras11; CMNF17]. In the following proposition, we show that even a relaxed version of this pointwise comparison must necessarily yield a bound scaling as  $d^{\alpha-1}$ . In contrast, we show that the comparison between the *maximum* difference of the two entropies exhibits much better scaling, and in fact can yield the Lipschitz continuity bound of (5.17). This can be understood by the fact that the Tsallis entropy is a Concave-Type  $(h, \phi)$ -entropy, and hence it increases the slowest near the uniform distribution (in the sense that Theorem 5.1.1 holds). On the other hand, for  $\alpha > 1$ , the  $\alpha$ -Rényi entropy is a Convex-Type  $(h, \phi)$ -entropy, and increases quickest at a distribution of the form  $(x, z, \dots, z, y)$  for  $x \leq z \leq y$  as shown by Theorem 5.1.4, which can be close to uniform.

Recall the definition  $\Delta_\varepsilon^H(r) := \max_{q \in B_\varepsilon(r)} (H(q) - H(r))$  for a function  $H : \mathcal{P}_d \rightarrow \mathbb{R}$ .

**Proposition 5.2.5.** *Let  $\alpha > 1$ . The smallest constant  $c$  such that*

$$\Delta_\varepsilon^{S_\alpha}(r) \leq c \Delta_\varepsilon^{T_\alpha}(r) \quad (5.23)$$

for all  $r \in \mathcal{P}$  and  $\varepsilon > 0$  is  $c = \frac{d^{\alpha-1}}{\ln(2)}$ . However, the smallest constant  $\tilde{c}$  such that

$$\max_{r \in \mathcal{P}} \Delta_\varepsilon^{S_\alpha}(r) \leq \tilde{c} \max_{r' \in \mathcal{P}} \Delta_\varepsilon^{T_\alpha}(r') \quad (5.24)$$

for all  $\varepsilon > 0$  satisfies  $\tilde{c} \leq \frac{\alpha d}{\ln(2)}$ .

*Remark.* (5.23) shows that at some points  $r \in \mathcal{P}$  the Tsallis entropy can locally increase much more than the Rényi entropy, while (5.24) shows that their maximal local increases are not so different.

*Proof.* Since

$$\Delta_\varepsilon^{S_\alpha}(r) = \int_0^\varepsilon \Gamma_{S_\alpha}(\mathcal{M}_\varepsilon(r)) \, ds$$



and

$$\Delta_\varepsilon^{T_\alpha}(r) = \int_0^\varepsilon \Gamma_{T_\alpha}(\mathcal{M}_\varepsilon(r)) \, ds$$

it suffices to bound the ratio  $\frac{\Gamma_{S_\alpha}(p)}{\Gamma_{T_\alpha}(p)}$  uniformly in  $p \in \mathcal{P}$  by  $c$ . On the other hand, for (5.24) to hold for all  $\varepsilon > 0$  and  $r \in \mathcal{P}$ , the same ratio must in fact be bounded by  $c$ . We have

$$\Gamma_{S_\alpha}(p) = \frac{1}{\ln(2)} \frac{\alpha}{\alpha - 1} \frac{r_+^{\alpha-1} - r_-^{\alpha-1}}{\sum_{i=1}^d p_i^\alpha}, \quad \Gamma_{T_\alpha}(p) = \frac{\alpha}{\alpha - 1} (r_+^{\alpha-1} - r_-^{\alpha-1})$$

and hence

$$\frac{\Gamma_{S_\alpha}(p)}{\Gamma_{T_\alpha}(p)} = \frac{1}{\ln(2)} \left( \sum_{i=1}^d p_i^\alpha \right)^{-1}.$$

Since  $p \mapsto \sum_{i=1}^d p_i^\alpha$  is Schur convex, the above ratio is Schur concave, and hence maximized at the uniform distribution. Thus,

$$\max_{p \in \mathcal{P}} \frac{\Gamma_{S_\alpha}(p)}{\Gamma_{T_\alpha}(p)} = \frac{1}{\ln(2)} \left( d \left( \frac{1}{d} \right)^\alpha \right)^{-1} = \frac{1}{\ln(2)} d^{\alpha-1}.$$

To estimate  $\tilde{c}$ , we simply rewrite the uniform continuity bound for  $T_\alpha$  given in Equation (5.6) as:

$$\max_{r' \in \mathcal{P}} \Delta_\varepsilon^{T_\alpha}(r') = \frac{1}{\alpha - 1} (1 - (1 - \varepsilon)^\alpha - (d - 1)^{1-\alpha} \varepsilon^\alpha),$$

noting that the maximum is achieved at  $r' = (1, 0, \dots, 0)$ . We have  $(1 - \varepsilon)^\alpha \leq 1 - \varepsilon$  and hence  $1 - (1 - \varepsilon)^\alpha \geq \varepsilon$ . Then

$$\max_{r' \in \mathcal{P}} \Delta_\varepsilon^{T_\alpha}(r') \geq \frac{\varepsilon}{\alpha - 1}.$$

On the other hand, (5.17) yields

$$\max_{r \in \mathcal{P}} \Delta_\varepsilon^{S_\alpha}(r) \leq \varepsilon \frac{\alpha}{\alpha - 1} \frac{d}{\ln(2)} \quad \square$$

Proposition 5.2.5 helps show that the notion of majorization flow can provide additional insight into the entropy landscape with respect to the TV distance, beyond the entropic bounds it can establish.

### 5.2.3 Connection to thermodynamics

[Bae11] introduced an interesting connection between the  $\alpha$ -Rényi entropy and free energies in thermodynamics. In this section, we recall this relationship, and remark on

the resulting consequences of our continuity bounds for Rényi entropies of Gibbs states and their connection to changes in free energy.

The following holds in either a quantum or classical picture. We will work in quantum notation for consistency with [Bae11]. Consider a Gibbs state

$$\rho(T) = Z(T)^{-1} e^{-H/T}$$

where  $H$  is the Hamiltonian,  $T \geq 0$  the temperature,  $Z(T) = \text{tr}(e^{-H/T})$  is the partition function, and we have set Boltzmann's constant  $k_B \equiv 1$ . We can define the free energy as

$$F(T) = -T \ln Z(T).$$

By direct calculation, we find that the  $\alpha$ -Rényi entropy  $S_\alpha$  satisfies

$$S_{\frac{T_0}{T}}(\rho(T_0)) = -\frac{F(T) - F(T_0)}{T - T_0}. \quad (5.25)$$

for any  $T > 0$  [Bae11, Equation (9)]. In the limit  $T \rightarrow T_0$ , we recover the thermodynamic relation

$$S(\rho(T_0)) = -\left. \frac{dF}{dT} \right|_{T_0}$$

that the entropy is the derivative of the free energy with respect to temperature. Note that any full-rank state  $\sigma$  can be seen as a Gibbs state at temperature  $T$  associated to the Hamiltonian  $H = -\frac{1}{T} \log \sigma$ . This gives a physical interpretation to  $S_\alpha(\sigma)$  for any full-rank state  $\sigma$ : consider  $\sigma$  to be the Gibbs state at initial temperature  $T_0 = 1$ . Then  $S_\alpha(\sigma)$  equals the negative of the ratio of the change in free energy to the change in temperature when the temperature is changed from  $T_0$  to  $\alpha^{-1}T_0$ . This can be seen as the maximum amount of work the system, initially in thermal equilibrium at temperature  $T_0$ , can do when its temperature is suddenly changed from  $T_0$  to  $\alpha^{-1}T_0$  as it moves to the new thermal equilibrium, divided by the change in temperature [Bae11].

The Schur concavity of  $S_\alpha$  for all  $\alpha > 0$  can be interpreted through this physical picture as well. The relation  $\rho \prec \sigma$  means that the distribution of the eigenvalues of  $\rho$  is “flatter” and “more disordered” than those of  $\sigma$ ; correspondingly,  $H_\rho := -\frac{1}{T} \log \rho$  log-majorizes  $H_\sigma := -\frac{1}{T} \log \sigma$ , where log majorization is defined by  $A \prec_{\log} B$  if  $\log A \prec \log B$ . By Schur concavity, if  $\rho \prec \sigma$ , then  $S_\alpha(\rho) \geq S_\alpha(\sigma)$  for any  $\alpha > 0$ . Hence, the Schur concavity of the  $\alpha$ -Rényi entropy can be interpreted as a statement about how the distribution of energy levels of a Hamiltonian relates to the free-energy increase or decrease of the system (per unit change in temperature) under a sudden change in temperature.

Now, recall that the trace distance is endowed with an operational interpretation in terms of distinguishability under measurement. We say that  $\rho$  and  $\sigma$  are  $\varepsilon$ -indistinguishable if  $\frac{1}{2}\|\rho - \sigma\|_1 \leq \varepsilon$ . Consider an experiment in which the system is in a state  $\sigma$  which is not known precisely, but is  $\varepsilon$ -indistinguishable from a known state  $\rho$ , which is a Gibbs state,  $\rho = \rho(T_0)$ . We pose the following question:

If the temperature is abruptly changed from  $T_0 \rightarrow T$ , can one bound the ratio of the change in free energy and the corresponding change in temperature?

This quantity is exactly the  $\alpha$ -Rényi entropy of order  $\alpha = T_0/T$  by (5.25), and hence (5.16) and Proposition 5.2.2 provide an answer in the affirmative. Moreover, Proposition 5.2.2 shows that the  $\alpha$ -Rényi entropy is Lipschitz continuous if and only if  $\alpha > 1$ . In other words, if  $T_0 > T$ , then there exists  $k_{T_0/T} < \infty$  such that

$$|S_{T_0/T}(\sigma) - S_{T_0/T}(\rho(T_0))| \leq \varepsilon k_{T_0/T}.$$

If  $T_0 \leq T$ , then no such linear bound can hold uniformly in  $\sigma$  and  $\rho(T_0)$ , but (5.6) gives a tight uniform (nonlinear) bound.

## 5.3 Other entropies

- The von Neumann entropy

$$S(\rho) = -\operatorname{tr}(\rho \log \rho).$$

satisfies  $S = S_{(h,\phi)}$  for  $h(x) = x$  and  $\phi(x) = -x \log x$ , and is a Concave-Type  $(h, \phi)$ -entropy. The Audenaert-Fannes bound, (1.1), provides a tight uniform continuity bound for the von Neumann (or Shannon) entropy.

**Our contribution** Corollary 5.1.2 provides an alternate proof for (1.1) and the necessary and sufficient conditions for equality as a consequence of the fact that the von Neumann entropy is a Concave-Type  $(h, \phi)$ -entropy, and Corollary 5.1.3 shows that  $S$  is not Lipschitz continuous with respect to the trace distance, but its smoothed variant  $S^\delta$  is Lipschitz continuous for any  $\delta \in (0, 1)$ , with an optimal Lipschitz constant of

$$\log(\delta^{-1} - 1) + \log(d - 1). \quad (5.26)$$

- The  $(s, \alpha)$ -unified entropies,

$$E_\alpha^s(\rho) = \frac{1}{s(1-\alpha)}(\text{tr}[\rho^\alpha]^s - 1)$$

for  $\alpha \in (0, 1) \cup (1, \infty)$ ,  $s \in \mathbb{R} \setminus \{0\}$ , were introduced in the quantum case by [HY06] and in the classical case by [RT91]. This family of entropies includes the Tsallis entropies in the case  $s = 1$ , and the  $\alpha$ -Rényi entropies (up to a factor of  $\ln(2)$ ) in the limit  $s \rightarrow 0$ . We have  $E_\alpha^s = H_{(h, \phi)}$  for  $\phi(x) = x^\alpha$ , and  $h(x) = \frac{1}{s(1-\alpha)}(x^s - 1)$ , which satisfy  $\phi(0) = 0$ , and  $h(\phi(1)) = 0$ . If  $\alpha \in (0, 1)$ ,  $h$  is strictly increasing and  $\phi$  is strictly concave, while if  $\alpha > 1$ ,  $h$  is strictly decreasing, and  $\phi$  is strictly convex. Additionally,  $h$  is convex if  $(s > 1 \text{ and } \alpha < 1)$  or if  $(s < 1 \text{ and } \alpha > 1)$ , and is concave otherwise. Thus, if  $0 < \alpha < 1$  and  $s \leq 1$ ,  $E_\alpha^s$  is a Concave-Type  $(h, \phi)$ -entropy, and if  $\alpha > 1$  with  $s \leq 1$ , then  $E_\alpha^s$  is a Convex-Type  $(h, \phi)$ -entropy. If  $s > 1$ , then for any  $\alpha \in (0, 1) \cup (1, \infty)$ ,  $E_\alpha^s$  is an  $(h, \phi)$ -entropy in the sense defined by [SMMP93], but not of Concave-Type or Convex-Type, and hence the results of Section 5.1 do not apply in that case.

*Remark.* [HY06, Proposition 5] incorrectly claims that the unified entropies are not Schur concave. However, they are indeed strictly Schur concave for all  $\alpha \in (0, 1) \cup (1, \infty)$  and  $s \in \mathbb{R} \setminus \{0\}$ .

Rastegin [Ras11] showed that for  $0 < \alpha < 1$ ,  $s \in (-\infty, -1] \cup [0, 1]$ , and  $T(\rho, \sigma) \leq \frac{1}{2}\alpha^{\frac{1}{1-\alpha}}$ , the bound

$$|E_\alpha^s(\rho) - E_\alpha^s(\sigma)| \leq (2\varepsilon)^\alpha \ln_\alpha d + n_\alpha(2\varepsilon) \quad (5.27)$$

holds, where  $\ln_\alpha = \frac{x^{1-\alpha}-1}{1-\alpha}$  and  $n_\alpha(x) = \frac{x^\alpha-x}{1-\alpha}$ . If  $\alpha > 1$  and  $s \in [-1, 0] \cup [1, +\infty]$ ,

$$|E_\alpha^s(\rho) - E_\alpha^s(\sigma)| \leq \chi_s[\varepsilon^\alpha \ln_\alpha(d-1) + t_\alpha(\varepsilon)] \quad (5.28)$$

where  $t_\alpha(\varepsilon) := T_\alpha(\{\varepsilon, 1-\varepsilon\})$  is the binary Tsallis entropy. In [HY06, Proposition 6], the Lipschitz continuity bound

$$|E_\alpha^s(\rho) - E_\alpha^s(\sigma)| \leq \frac{\alpha}{\alpha-1} \varepsilon \quad (5.29)$$

for any  $\rho, \sigma \in \mathcal{D}(\mathcal{H})$  with  $T(\rho, \sigma) \leq \varepsilon$  for  $\alpha > 1$ , and  $s \geq 1$  was proven.

**Our contribution:**

For  $\alpha \in (0, 1)$  and  $s \leq 1$ , then  $E_\alpha^s$  is not Lipschitz continuous on  $\mathcal{D}(\mathcal{H})$  by Corollary 5.1.3, but satisfies the following tight uniform continuity bound by Corollary 5.1.2: If  $\varepsilon \in [0, 1]$  and  $\rho, \sigma \in \mathcal{D}$  with  $T(\rho, \sigma) \leq \varepsilon$ ,

$$|E_\alpha^s(\rho) - E_\alpha^s(\sigma)| \leq \begin{cases} \frac{1}{s(1-\alpha)} [((1-\varepsilon)^\alpha + (d-1)^{1-\alpha}\varepsilon^\alpha)^s - 1] & \varepsilon < 1 - \frac{1}{d} \\ \frac{1}{s(1-\alpha)} [d^{s(1-\alpha)} - 1] & \varepsilon \geq 1 - \frac{1}{d} \end{cases} \quad (5.30)$$

with equality if and only if one state is pure, and the other state has spectrum  $\{1 - \varepsilon, \frac{\varepsilon}{d-1}, \dots, \frac{\varepsilon}{d-1}\}$  if  $\varepsilon < 1 - \frac{1}{d}$ , or  $\{\frac{1}{d}, \dots, \frac{1}{d}\}$  if  $\varepsilon \geq 1 - \frac{1}{d}$ .

If  $\alpha > 1$  and  $s \leq 1$ , then  $E_\alpha^s$  is Lipschitz continuous on  $\mathcal{D}(\mathcal{H})$  by Theorem 5.1.4, and the associated optimal Lipschitz constant  $k_\alpha^s$  satisfies

$$k_\alpha^s \leq \begin{cases} \frac{\alpha}{\alpha-1} d^{1-\alpha s} & s\alpha < 1 \\ \frac{\alpha}{\alpha-1} & s\alpha \geq 1. \end{cases} \quad (5.31)$$

- Entropies induced by divergences. Denoting left multiplication by an operator  $A$  as  $L_A$ , and right multiplication by  $A$  as  $R_A$ , one defines the  $f$ -divergence

$$S_f(\rho||\sigma) := \text{tr}[\sigma^{1/2} f(L_\rho R_{\sigma^{-1}})(\sigma^{1/2})] \quad (5.32)$$

which was first introduced by Petz [Pet85] (see [HM17] for a useful overview). The *maximal  $f$ -divergence* [PR98] is given by

$$\hat{S}_f(\rho||\sigma) := \text{tr}[\sigma^{1/2} f(\sigma^{-1/2} \rho \sigma^{-1/2}) \sigma^{1/2}]. \quad (5.33)$$

From either divergence, one can define an associated entropy by evaluating at  $\sigma = \mathbf{1}$  (and reversing the sign). The two entropies coincide, yielding

$$S_f(\rho) := -S_f(\rho||\mathbf{1}) = -\hat{S}_f(\rho||\mathbf{1}) = -\text{tr}[f(\rho)]. \quad (5.34)$$

For strictly convex  $f$  with  $f(0) = f(1) = 0$  we can define  $\phi = -f$  and  $h(x) = x$ , yielding a Concave-Type  $(h, \phi)$ -entropy.

**Our contribution:** For strictly convex  $f$  with  $f(0) = f(1) = 0$ , Corollary 5.1.2 gives that for  $\varepsilon \in [0, 1]$  and  $\rho, \sigma \in \mathcal{D}(\mathcal{H})$  with  $T(\rho, \sigma) \leq \varepsilon$ ,

$$|S_f(\rho) - S_f(\sigma)| \leq \begin{cases} -f(1 - \varepsilon) - (d - 1)f(\frac{\varepsilon}{d-1}) & \varepsilon < 1 - \frac{1}{d} \\ -df(\frac{1}{d}) & \varepsilon \geq 1 - \frac{1}{d}. \end{cases} \quad (5.35)$$

Corollary 5.1.3 shows that  $S_f$  is Lipschitz continuous on  $\mathcal{D}(\mathcal{H})$  if and only if

$$k := \lim_{\varepsilon \rightarrow 0} -\frac{1}{\varepsilon}f(1 - \varepsilon) - (d - 1)\frac{1}{\varepsilon}f(\frac{\varepsilon}{d-1}) \quad (5.36)$$

is finite. In the latter case,  $k$  is the optimal Lipschitz constant for  $S_f$ . Note that if  $f$  is differentiable at 0 and 1, then  $k = f'(1) - f'(0)$ .

- The *concurrence* of a bipartite pure state  $\psi_{AB}$  is an entanglement monotone defined as

$$C(\psi_{AB}) = \sqrt{2(1 - \text{tr}[\rho_A^2])}$$

where  $\rho^\psi = \text{tr}_B[\psi_{AB}]$  is the reduced state on system  $A$  [Woo01; RBCH+01]. Regarded as a function of the reduced state, the concurrence can be seen as Concave-Type  $(h, \phi)$ -entropy with  $\phi(x) = -x^2$  and  $h(x) = \sqrt{2(1 + x)}$ , and hence Corollary 5.1.2 gives a tight uniform continuity bound in terms of the trace distance between the reduced states. If for some  $\varepsilon > 0$  two bipartite pure states  $\psi_{AB}$  and  $\phi_{AB}$  satisfy  $T(\psi_{AB}, \phi_{AB}) \leq \varepsilon$ , then by monotonicity of the trace distance under partial trace,  $T(\rho^\psi, \rho^\phi) \leq \varepsilon$  as well. Hence, Corollary 5.1.2 yields

$$|C(\psi_{AB}) - C(\phi_{AB})| \leq \begin{cases} \sqrt{2(1 - (1 - \varepsilon)^2 - (d - 1)^{-1}\varepsilon^2)} & \varepsilon < 1 - \frac{1}{d} \\ \sqrt{2(1 - d^{-1})} & \varepsilon \geq 1 - \frac{1}{d} \end{cases} \quad (5.37)$$

for all bipartite pure states  $\psi_{AB}$  and  $\phi_{AB}$  such that  $T(\psi_{AB}, \phi_{AB}) \leq \varepsilon$ . The concurrence is not Lipschitz continuous, by Corollary 5.1.3.

## 5.4 Number of distinct realizations from $N$ i.i.d. random variables

In this section, we consider another application of the majorization flow to continuity bounds.

Consider an experiment in which outcome  $i \in \{1, \dots, M\}$  is observed with probability  $p_i$ , for some probability distribution  $p \in \mathcal{P}_M$ . Repeat this experiment  $N$  times, independently, and consider the random variable  $K$  which denotes the number of distinct outcomes observed. In the following, we establish the Lipschitz continuity of each entry of the cumulative probability distribution (c.d.f.) of  $K$ , as a function of  $p$ , a tight uniform continuity bound for  $\mathbb{E}[K]$ , and the optimal Lipschitz constant for  $\mathbb{E}[K]$ , as functions of  $p$ . These results quickly follow from the results of Chapter 4 and the expressions for the c.d.f. of  $K$  and  $\mathbb{E}[K]$  given in [WY73]. These bounds provide a notion of robustness in order to understand how much  $K$  can vary given some uncertainty in  $p$ .

This experiment was introduced in the excellent review [Arn07] as describing an experiment on an island in which there is some unknown (but fixed) number  $M$  of butterfly species, and butterflies are sequentially trapped until  $N$  individuals have been caught. In this context,  $K$  is the number of distinct species of butterflies observed, and  $p$  is an underlying “catchability” distribution which gives the probability  $p_j > 0$  that a butterfly of species  $j$  will be caught on any given trial<sup>4</sup>. In this scenario, the continuity bounds for  $K$  provide a notion of robustness by describing how much the number of distinct observed species can vary given some change in the catchability distribution.

In [WY73] it was shown that each entry of the cumulative distribution of  $K$ , namely

$$F_j \equiv F_j(p) := \Pr[K \leq j],$$

is a Schur convex function of  $p$  (Theorem 4.1 of the above article), and has the expression

$$F_j = \sum_{i=1}^j (-1)^{j-i} \binom{M-i-1}{j-i} \sum_{1 \leq l_1 < l_2 < \dots < l_i \leq M} (p_{l_1} + \dots + p_{l_i})^N$$

which is given in [WY73, Corollary 3.2]. Taking the derivative, we obtain

$$\frac{\partial F_j}{\partial p_1} - \frac{\partial F_j}{\partial p_2} = f_j(p_1) - f_j(p_2)$$

for

$$f_j(s) = N \sum_{i=1}^j (-1)^{j-i} \binom{M-i-1}{j-i} \sum_{\substack{1 < l_2 < \dots < l_i \leq M \\ l_2, \dots, l_i \neq 1, 2}} (s + p_{l_2} + \dots + p_{l_i})^{N-1}$$

---

<sup>4</sup>The independence of the trials corresponds to an assumption that there is a large number of individuals of each species such that trapping one does not influence the probability of which species will be caught on the next trial.

as was calculated in [WY73, Equation 11]. To compute the optimal Lipschitz constant for  $F_j$ , it remains to maximize this difference over  $p \in \mathcal{P}_M$ . We leave that for future work, and simply show that  $F_j$  is a Lipschitz continuous function of  $p \in \mathcal{P}_M$  by showing that  $f_j(p_1) - f_j(p_2) < \infty$  for any  $p \in \mathcal{P}_M$ . We can use that the summand  $(s + p_{l_2} + \cdots + p_{l_i})$  is less than 1 for  $s \in \{p_1, p_2\}$ , and that the number of elements in the second summation is  $\binom{M-2}{i}$  to find the simple bound

$$|f_j(s)| \leq N \sum_{i=1}^j \binom{M-i-1}{j-i} \binom{M-2}{i} < \infty$$

for  $s \in \{p_1, p_2\}$ , which completes the proof.

Next, [WY73, Corollary 3.3] shows that the expected number of distinct elements,  $\mathbb{E}[K]$  satisfies

$$\mathbb{E}[K] \equiv \mathbb{E}_p[K] = M - \sum_{i=1}^M (1 - p_i)^N.$$

In fact, we can identify  $\mathbb{E}[K] - 1$  as a Concave-Type  $(h, \phi)$ -entropy, with  $h(x) = x - 1$  and  $\phi(x) = 1 - (1 - x)^N$ , as defined in Section 5.1. Hence, Corollary 5.1.2 shows that for  $\varepsilon > 0$  if  $p, q \in \mathcal{P}_M$  satisfy  $\text{TV}(p, q) \leq \varepsilon$ , then

$$|\mathbb{E}_p[K] - \mathbb{E}_q[K]| \leq \begin{cases} \frac{(M-1)^N - (M-1-\varepsilon)^N}{(M-1)^{N-1}} - \varepsilon^N & \varepsilon \leq 1 - \frac{1}{M} \\ \frac{M^N - (M-1)^N}{M^{N-1}} - 1 & \varepsilon > 1 - \frac{1}{M}. \end{cases}$$

In particular, using Corollary 5.1.3,

$$|\mathbb{E}_p[K] - \mathbb{E}_q[K]| \leq \varepsilon N, \tag{5.38}$$

and  $N$  is the optimal Lipschitz constant. Note that (5.38) does not depend on  $M$ , but its derivation assumes  $M < \infty$ .

## 5.5 Number of connected components of a random graph

In this section, we examine the continuity of a quantity with a fairly different flavor: the expected number of connected components  $E_C(p)$  of a particular random graph model which is parametrized by a probability distribution  $p \in \mathcal{P}$ . This quantity has the two necessary features to apply the tools of Section 4.4: it is a Schur concave function



of a probability vector, and it has a closed-form expression which we can differentiate. Moreover, the bounds we obtain are quite sharp: we obtain the optimal Lipschitz constant, which scales as  $\sqrt{n-2}$  for  $n$ -dimensional distributions, up to an additive factor of at most five and a multiplicative factor of at most  $\sqrt{2}$ , for all  $n \geq 5$ .

We consider the following random graph  $G$  with  $n \in \mathbb{N}$  nodes (or vertices), which was constructed in [Ros81]. Fix a probability distribution  $p$  on  $\{1, \dots, n\}$ , and take  $n$  independent and identically distributed (i.i.d.) random variables  $X_1, \dots, X_n \sim p$  such that  $\Pr[X_i = j] = p_j$  for all  $i, j \in \{1, 2, \dots, n\}$ . Then construct  $n$  edges by connecting  $i$  to  $X_i$ . The result is a graph with  $n$  nodes and edges, such that every node has at least one edge.

**Definition 5.5.1.** A connected component of a graph  $G$  is a subgraph  $H$  such that for every pair of nodes  $x, y \in H$ , there is a path (made up of contiguous edges) between  $x$  and  $y$ , and moreover there are no edges between nodes in  $H$  and  $G \setminus H$  (in other words, no edges leave the connected component).

As a simple example, if  $p = (1, 0, \dots, 0)$ , then 1 has a single self-edge, and every other node has a single edge touching it, which is connected to 1, and hence the graph has one connected component. As another example, if  $p = (1/3, 0, 2/3, 0, \dots, 0)$ , then one realization of the random graph is shown in Figure 5.1, and another shown in Figure 5.3.

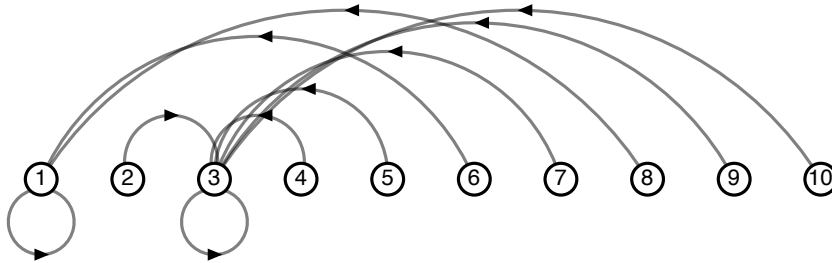


Fig. 5.1 A realization of the random graph  $G$  associated to the distribution  $p = (1/3, 0, 2/3, 0, \dots, 0)$ . A set of edges  $\{(i, X_i) : i = 1, \dots, 10\}$  is constructed by independently sampling each  $X_i \sim p$ . The edges shown here are directed, but for the purpose of calculating the number of connected components, we consider the associated undirected graph.

The number of connected components  $C$  of  $G$ , the random graph constructed above, is a random variable. Its expectation satisfies

$$\mathbb{E}[C] = \sum_{\substack{S \subseteq \{1, \dots, n\} \\ S \text{ nonempty}}} (|S| - 1)! \prod_{j \in S} p_j \quad (5.39)$$

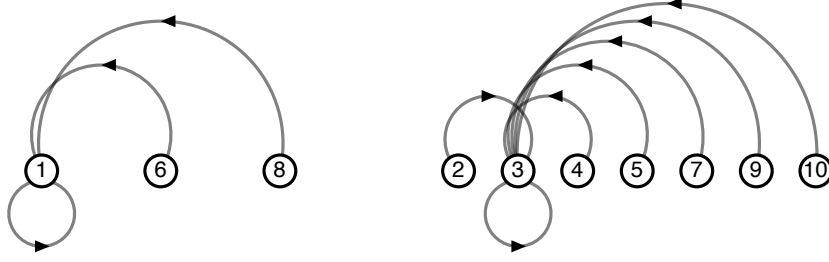


Fig. 5.2 The two connected components of the same realization of the graph  $G$  from Figure 5.1. These consist of elements which connect to  $\{1\}$ , and elements which connect to  $\{3\}$ , respectively. Note that we neglect directionality in computing connected components.

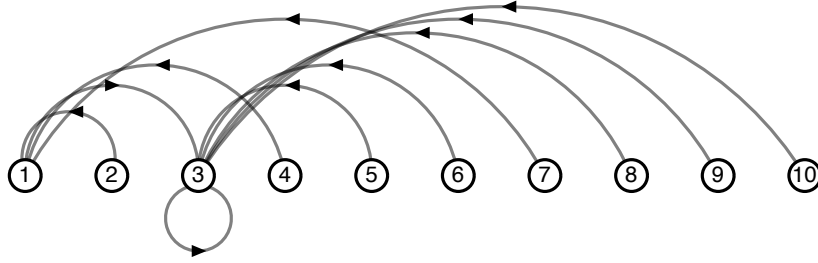


Fig. 5.3 Another realization of the graph  $G$  from Figure 5.1, with underlying probability distribution  $p = (1/3, 0, 2/3, 0, \dots, 0)$ . This realization has only one connected component.

by [Ros81, Equation 3]. To make the dependence on  $p$  explicit, let us write

$$E_C(p) \equiv \mathbb{E}[C]. \quad (5.40)$$

Note that  $E_C(p)$  is a polynomial in the entries of  $p$ , and hence is smooth. On the compact set  $\mathcal{P}_n$ ,  $E_C$  is thus a Lipschitz function. In the following, we establish upper and lower bounds on the optimal Lipschitz constant. These bounds involve the Mills ratio,

$$M(x) := \frac{1 - \Phi(x)}{\phi(x)} \quad (5.41)$$

where  $\phi$  and  $\Phi$  are the probability mass function and cumulative density function, respectively, of the standard normal distribution.

**Theorem 5.5.2.** *Let  $E_C(p)$  and  $E_C(q)$  denote the expected number of connected components of the random graph of  $n \geq 3$  nodes, corresponding to probability distributions  $p$*

and  $q$ , respectively. Let  $f(x) := x - x^2 M(x)$ , and  $\mu$  be its maximum value on the domain  $x \geq 0$ . Then the following bound holds:

$$|E_C(p) - E_C(q)| \leq (3 + \mu\sqrt{n-2}) \text{TV}(p, q). \quad (5.42)$$

Moreover, setting  $x_0$  to be the unique maximizer of  $f$  on  $x \geq 0$ , we have that any Lipschitz constant  $\kappa$  satisfies

$$\kappa \geq \frac{\mu\sqrt{n-2}}{\sqrt{2}} - \frac{\mu x_0}{2} - \sqrt{\frac{2}{n-2}} x_0 - \frac{x_0^2}{n-2} - \sqrt{n-2} e^{-(n-2)} \frac{x_0 e^{x_0^2/2}}{\sqrt{2}}. \quad (5.43)$$

Additionally,  $x_0$  and  $\mu$  satisfy the following explicit bounds<sup>5</sup>:

$$1.1615278892744612 \leq x_0 \leq 1.1615278892744958, \quad (5.44)$$

$$0.346813047097384 \leq \mu \leq 0.346813047097549. \quad (5.45)$$

*Remark.* For  $n \geq 5$ , the lower order terms in (5.43) satisfy

$$-\frac{\mu x_0}{2} - \sqrt{\frac{2}{n-2}} x_0 - \frac{x_0^2}{n-2} - \sqrt{n-2} e^{-(n-2)} \frac{x_0 e^{x_0^2/2}}{\sqrt{2}} \geq -2$$

yielding the simpler lower bound  $\kappa \geq \frac{\mu\sqrt{n-2}}{\sqrt{2}} - 2$  for  $n \geq 5$ .

**Connection to models of epidemics** This kind of random graph model has been used for modelling epidemics; in fact,  $G$  corresponds to the “two-sided epidemic process” of [Ger77]. We can interpret the continuity bound given in Theorem 5.5.2 in this framework as follows. Let the nodes of the graph represent people (or, e.g., groups of people), and let the edges represent pairwise interactions between people which can spread the infection. We can see the probability distribution  $p$  governing the placement of edges as being associated to a model of interactions. Now, what is the minimum number of people who need to be infected initially in order for the whole population (all  $n$  nodes) to become infected eventually? The answer is simply the number of connected components of the graph. This is because if one person per connected component is infected, that person can then spread the infection to the rest of the people in the connected component. On the other hand, any connected component lacking an infected person will never become infected.

---

<sup>5</sup>These bounds can be refined to arbitrary precision as shown in [Han20a].

In this interpretation, (5.39) provides a formula for computing  $E_C(p)$ , the expected minimal number of initially infected people required to infect the whole population. However, if interactions are better modelled by some  $p' \neq p$ , then  $E_C(p')$  provides a better estimate of the minimal number of people required to infect the whole population. Theorem 5.5.2 then provides a quantitative measure of the *robustness* of  $E_C$  by establishing a bound on the error  $|E_C(p) - E_C(p')|$  as a function of the model error given by the total variation distance between  $p$  and  $p'$ . There are, of course, many other measures of robustness that may be useful in this setting which are not considered here; e.g. of the average against fluctuations, of the framework of the model itself, and so forth.

**Overview of the remainder of the section** The upper bound (5.42) is proven in Section 5.5.1, and the lower bound (5.43) is proven in Section 5.5.2. The uniqueness of the maximizer of  $f(x) = x - x^2 M(x)$  and the bounds on the maximizer  $x_0$  and the maximum value  $\mu$  are established using interval arithmetic, as described in Section 5.5.3.

### 5.5.1 Proof of the upper bound (5.42)

In this section, we use Corollary 4.4.1 to establish Theorem 5.5.2. Note that by (5.39),  $E_C : \mathcal{P}_n \rightarrow \mathbb{R}$  is a polynomial in the components of the probability vector  $p$  and in particular is continuously differentiable. In [Ros81, Proposition 1], the author proves that  $p \mapsto E_C(p)$  is Schur concave using the criterion (2.5), by showing that

$$\partial_{p_i} E_C(p) - \partial_{p_j} E_C(p) = (p_j - p_i) \sum_{S^*} (|S^*| + 1)! \prod_{j \in S^*} p_j$$

where  $S^*$  ranges over nonempty sets of  $\{1, \dots, n\} \setminus \{i, j\}$ . Hence,

$$\Gamma_{E_C}(r) = (r_+ - r_-) \sum_{S^*} (|S^*| + 1)! \prod_{j \in S^*} r_j$$

where  $S^*$  ranges over nonempty sets of  $I := \{1, \dots, n\} \setminus \{i_+, i_-\}$ , where  $i_{\pm}$  are indices such that  $r_{i_{\pm}} = r_{\pm}$ . We can use the criterion (2.5) again by repeating the proof of [Ros81, Proposition 1] to show that for

$$S(\{r_i\}_{i \in I}) := \sum_{S^*} (|S^*| + 1)! \prod_{j \in S^*} r_j,$$

we have

$$\partial_{r_i} S(\{r_i\}_{i \in I}) - \partial_{r_j} S(\{r_i\}_{i \in I}) = (r_j - r_i) \sum_{S^*} (|S^*| + 3)! \prod_{j \in S^*} r_j$$

and hence  $S$  is Schur concave on the set  $\{p \in \mathbb{R}^{n-2} : p_i \geq 0, \sum_i p_i = 1 - r_+ - r_i\}$ . For such  $p$ ,

$$S(p) \leq S\left(\left\{\frac{1 - r_- - r_+}{n - 2}\right\}_{i \in I}\right)$$

and thus

$$\Gamma_{E_C}(r) \leq (r_+ - r_-) \sum_{k=1}^{n-2} \binom{n-2}{k} (k+1)! (1 - r_- - r_+)^k (n-2)^{-k}. \quad (5.46)$$

To obtain a Lipschitz bound on  $E_C(r)$ , it suffices to bound  $\Gamma_{E_C}(r)$  independently of  $r \in \mathcal{P}$ . We upper bound (5.46) by taking  $r_- = 0$ . For the simplicity of notation, let  $s = r_+$  and  $m = n - 2$ . Then we aim to bound

$$B_m(s) := s \sum_{k=1}^m \binom{m}{k} (k+1)! (1-s)^k m^{-k} \quad (5.47)$$

for  $s \in [\frac{1}{m+2}, 1]$ , using that  $r_+ \in [\frac{1}{n}, 1]$  which follows from  $r \in \mathcal{P}$ . Let

$$S_m(s) := \sum_{k=1}^m c_{k,m} (1-s)^{k-1}$$

with

$$c_{k,m} := \binom{m}{k} \frac{(k+1)!}{m^k} = (k+1) \prod_{j=1}^{k-1} \left(1 - \frac{j}{m}\right),$$

then

$$B_m(s) = s(1-s)S_m(s). \quad (5.48)$$

Applying the inequality  $1 - x \leq \exp(-x)$  to every factor in  $c_{k,m}$  gives the simple upper bound

$$\begin{aligned} c_{k,m} &\leq (k+1) \prod_{j=1}^{k-1} \exp(-j/m) \\ &= (k+1) \exp\left(-\sum_{j=1}^{k-1} j/m\right) = (k+1) \exp\left(-\frac{(k-1)k}{2m}\right) \\ &\leq (k+1) \exp\left(-\frac{(k-1)^2}{2m}\right). \end{aligned}$$

As  $c_{k,m} \geq 0$ , we can also use the same inequality  $1 - s \leq \exp(-s)$  in the formula for  $S_m(s)$ . This gives as a first upper bound:

$$\begin{aligned} S_m(s) &\leq \sum_{k=1}^m (k+1) \exp\left(-\frac{(k-1)^2}{2m} - (k-1)s\right) \\ &= \sum_{l=0}^{m-1} (l+2) \exp\left(-\frac{l^2}{2m} - ls\right) \\ &= 2 + \sum_{l=1}^{m-1} (l+2) \exp\left(-\frac{l^2}{2m} - ls\right). \end{aligned}$$

We can interpret this sum as a lower Riemann sum for a certain Riemann integral. Noting that the factor  $l+2$  increases with  $l$  and the factor  $\exp\left(-\frac{l^2}{2m} - ls\right)$  decreases, we have

$$(l+2) \exp\left(-\frac{l^2}{2m} - ls\right) \leq \int_{l-1}^l (u+3) \exp\left(-\frac{u^2}{2m} - us\right) du.$$

Therefore,

$$\begin{aligned} S_m(s) &= 2 + \sum_{l=1}^{m-1} (l+2) \exp\left(-\frac{l^2}{2m} - ls\right) \\ &\leq 2 + \sum_{l=1}^{m-1} \int_{l-1}^l (u+3) \exp\left(-\frac{u^2}{2m} - us\right) du \\ &= 2 + \int_0^{m-1} (u+3) \exp\left(-\frac{u^2}{2m} - us\right) du \\ &\leq 2 + \int_0^{\infty} (u+3) \exp\left(-\frac{u^2}{2m} - us\right) du \\ &= 2 + \exp(ms^2/2) \int_0^{\infty} (u+3) \exp\left(-\frac{(u+ms)^2}{2m}\right) du. \end{aligned}$$

In terms of the probability density function  $\phi(x)$  of the standard normal distribution,  $\phi(x) = \exp(-x^2/2)/\sqrt{2\pi}$ , and making the substitution  $v = (u+ms)/\sqrt{m}$ , this last expression can be written as

$$\frac{1}{\phi(\sqrt{ms})} \int_0^{\infty} (u+3) \phi\left(\frac{u+ms}{\sqrt{m}}\right) du$$

$$\begin{aligned}
&= \frac{\sqrt{m}}{\phi(\sqrt{m}s)} \int_{\sqrt{m}s}^{\infty} (\sqrt{m}v - ms + 3)\phi(v) \, dv \\
&= \frac{\sqrt{m}}{\phi(\sqrt{m}s)} \left( \sqrt{m} \int_{\sqrt{m}s}^{\infty} v\phi(v) \, dv + (3 - ms) \int_{\sqrt{m}s}^{\infty} \phi(v) \, dv \right).
\end{aligned}$$

Exploiting the fact that  $x\phi(x) = -\phi'(x)$ , and with  $\Phi(x)$  the cumulative density function of the standard normal distribution, this last expression is equal to

$$\frac{\sqrt{m}}{\phi(\sqrt{m}s)} (\sqrt{m} \phi(\sqrt{m}s) + (3 - ms)(1 - \Phi(\sqrt{m}s))) = m + \sqrt{m}(3 - ms) \frac{1 - \Phi(\sqrt{m}s)}{\phi(\sqrt{m}s)},$$

so that

$$S_m(s) \leq 2 + m + \sqrt{m}(3 - ms) \frac{1 - \Phi(\sqrt{m}s)}{\phi(\sqrt{m}s)}.$$

The function in the last factor,

$$M(x) := \frac{1 - \Phi(x)}{\phi(x)},$$

is known as the Mills ratio, and several bounds are known for it. A well-known upper bound valid for  $x > 0$  is  $M(x) < 1/x$  [Gor41; YC15], which follows from the fact that  $M'(x) = xM(x) - 1$  and that  $M$  is a strictly decreasing function. Therefore,

$$3\sqrt{m}M(\sqrt{m}s) \leq 3/s,$$

and

$$S_m(s) \leq 2 + m + \frac{3}{s} - m^{3/2}sM(\sqrt{m}s).$$

Setting  $\mu$  to be as in Theorem 5.5.2, we have

$$-M(x) \leq \frac{\mu - x}{x^2}.$$

Therefore,

$$S_m(s) \leq 2 + m + \frac{3}{s} + \sqrt{m} \frac{\mu - \sqrt{m}s}{s} = 2 + \frac{3 + \mu\sqrt{m}}{s},$$

and

$$B_m(s) \leq (1 - s)(2s + 3 + \mu\sqrt{m}) = 2(1 - s)(1 + s) + (1 - s)(1 + \mu\sqrt{m}) \leq 2 + (1 + \mu\sqrt{m}),$$

over the interval  $0 \leq s \leq 1$ . □

Explicit numerical calculations of  $B_m(s)$  for  $m$  up to  $10^6$  suggest that the maximal value of  $B_m(s)$  is bounded below by  $\mu\sqrt{m}$  and, hence, lies within a constant not exceeding 3 of our bound, which is remarkable. In the following, we prove a slightly weaker bound, which recovers the square-root scaling at leading order.

### 5.5.2 Proof of the lower bound (5.43)

Let  $r = (r_+, \frac{1-r_+}{n-2}, \dots, \frac{1-r_+}{n-2}, 0) \in \mathcal{P}$  for some  $r_+ \in [\frac{1}{n-1}, 1]$ , so that  $r$  is a probability distribution with largest element  $r_+$ . Then the start of Section 5.5.1 establishes that

$$\Gamma_{E_C}(r) = B_m(s)$$

where  $B_m(s)$  is defined in (5.47), and  $s := r_+$ , and  $m := n - 2$ . By Corollary 4.4.1, it remains to lower bound  $B_m(s)$  for some  $s \in [\frac{1}{m+1}, 1]$ . As in (5.48), we write

$$B_m(s) = s \sum_{k=1}^m c_{k,m} (1-s)^k, \quad c_{k,m} := (k+1) \prod_{j=1}^{k-1} \left(1 - \frac{j}{m}\right). \quad (5.49)$$

Then

$$\ln \frac{c_{k,m}}{k+1} = \sum_{j=1}^{k-1} \ln \left(1 - \frac{j}{m}\right) = \sum_{j=0}^{k-1} \ln \left(1 - \frac{j}{m}\right) \geq \int_0^k \ln \left(1 - \frac{u}{m}\right) du$$

using that since  $j \mapsto \ln(1 - \frac{j}{m})$  is decreasing, the integral forms an underapproximation to the sum. By changing variables to  $v = u/m$ , we obtain

$$\ln \frac{c_{k,m}}{k+1} \geq m \int_0^{k/m} \ln(1-v) dv = -k - (m-k) \ln \left(1 - \frac{k}{m}\right) \geq -\frac{k^2}{m}$$

using that  $\ln(1 - \frac{k}{m}) \leq -\frac{k}{m}$ . Hence,

$$c_{k,m} \geq (k+1) \exp \left( -\frac{k^2}{m} \right). \quad (5.50)$$

From (5.49), defining  $c_{0,m} = 1$ , we have

$$\begin{aligned} \frac{1}{s} B_m(s) &= \sum_{k=1}^m c_{k,m} (1-s)^k = \sum_{k=0}^m c_{k,m} (1-s)^k - 1 \\ &\geq \sum_{k=0}^m (k+1) \exp \left( -\frac{k^2}{m} + k \ln(1-s) \right) - 1 \end{aligned}$$



$$\geq \int_0^{m+1} u \exp\left(-\frac{u^2}{m} + u \ln(1-s)\right) du - 1.$$

using (5.50) for the first inequality. For the second inequality, notice that the sum is of the form  $\sum_{k=0}^m f(k+1)g(k)$  where  $f(k) = k$  is monotone increasing, and  $g(k) = \exp\left(-\frac{k^2}{m} + k \ln(1-s)\right)$  is monotone decreasing. Hence, we have  $f(k+1) \geq \int_k^{k+1} f(u) du = \|f|_{[k,k+1]}\|_1$  and  $g(k) = \sup_{k \leq u \leq k+1} g(u) = \|g|_{[k,k+1]}\|_\infty$ , using that both functions are non-negative. Hölder's inequality gives

$$\int_k^{k+1} f(u)g(u) du \leq \|f|_{[k,k+1]}\|_1 \|g|_{[k,k+1]}\|_\infty \leq f(k+1)g(k)$$

and summing over  $k$  yields the inequality. Next, since

$$-\frac{u^2}{m} + u \ln(1-s) = -\frac{1}{m} \left( \left( u - \frac{m \ln(1-s)}{2} \right)^2 - \left( \frac{m \ln(1-s)}{2} \right)^2 \right),$$

we obtain

$$\begin{aligned} \frac{1}{s} B_m(s) &\geq \frac{1}{\exp\left(-\frac{1}{2} \left( \sqrt{\frac{m}{2}} \frac{\ln(1-s)}{2} \right)^2\right)} \int_0^{m+1} u \exp\left(-\frac{1}{2} \left( \sqrt{\frac{2}{m}} \left( u - \frac{m}{2} \ln(1-s) \right) \right)^2\right) du - 1 \\ &= \frac{1}{\phi(b)} \int_0^{m+1} u \phi(au - b) du - 1 \end{aligned}$$

for  $a = \sqrt{\frac{2}{m}}$ ,  $b = \sqrt{\frac{m}{2}} \ln(1-s)$ , and  $\phi(x) := \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$  is the p.d.f. of a standard normal distribution. Changing variables to  $v = au - b$ , we find

$$\begin{aligned} \frac{1}{s} B_m(s) &\geq \frac{1}{\phi(b)a^2} \left[ \int_{-b}^{a(m+1)-b} v \phi(v) dv + b \int_{-b}^{a(m+1)-b} \phi(v) dv \right] \\ &= \frac{1}{\phi(b)a^2} [\phi(-b) - \phi(a(m+1)-b) + b(\Phi(a(m+1)-b) - \Phi(-b))] - 1 \end{aligned}$$

where  $\Phi$  is the c.d.f of the standard normal distribution. Since  $a(m+1)-b \leq \sqrt{2m}$  and  $\phi(x)$  is decreasing on  $x > 0$ , we have  $\phi(a(m+1)-b) \leq \phi(\sqrt{2m})$ . Using also that  $\Phi(a(m+1)-b) \leq 1$ , we obtain

$$\frac{1}{s} B_m(s) \geq -\frac{\phi(\sqrt{2m})}{a^2 \phi(-b)} + \frac{1}{a^2} (1 + bM(-b)) - 1$$

where  $M(x) = \frac{1-\phi(x)}{\Phi(x)}$  is the Mills ratio. Substituting in  $a$ , we have

$$\frac{1}{s}B_m(s) \geq -\frac{m\phi(\sqrt{2m})}{2\phi(-b)} + \frac{m}{2}(1 + bM(-b)) - 1.$$

Recalling the definition of  $x_0$  and  $\mu$  from Theorem 5.5.2, we choose  $s = 1 - e^{-\sqrt{\frac{2}{m}}x_0}$  so that  $b = -x_0$ , and  $\mu = x_0 - x_0^2M(x_0)$ . Substituting for  $\mu$ , we have

$$B_m(s) \geq s \left( \frac{m\mu}{2x_0} - 1 - \frac{m\phi(\sqrt{2m})}{2\phi(x_0)} \right).$$

Using the bound  $e^{-x} \leq 1 - x + \frac{x^2}{2}$  for  $x \geq 0$ , we have  $s \geq \sqrt{\frac{2}{m}}x_0 - \frac{x_0^2}{m}$ . Hence,

$$B_m(s) \geq \frac{\mu\sqrt{m}}{\sqrt{2}} - \frac{\mu x_0}{2} - \sqrt{\frac{2}{m}}x_0 - \frac{x_0^2}{m} - \sqrt{m}e^{-m} \frac{x_0}{2\sqrt{\pi}\phi(x_0)}. \quad \square$$

### 5.5.3 Maximizing $x - x^2M(x)$

The function  $f(x) := x - x^2M(x)$  on the domain  $x \geq 0$  has a maximum value  $\mu$  satisfying (5.45) which occurs at a unique point  $x_0$  satisfying (5.44). To prove this, we will use the tools of *interval arithmetic*. Interval arithmetic is a method for rigorous calculation using finite-precision floating point numbers on a computer, as follows. Instead of considering a real number  $x \in \mathbb{R}$ , which may not be exactly representable with a particular finite precision arithmetic, a small interval  $[a, b] \subseteq \mathbb{R}$  containing  $x$  whose endpoints are exactly representable is used instead. Then to estimate e.g.  $f(x)$ , an interval  $[c, d] \subseteq \mathbb{R}$  is found such that  $f(y) \in [c, d]$  for any  $y \in [a, b]$ . This yields rigorous bounds on  $f(x)$  which are not subject to the “roundoff error” of usual floating point arithmetic. In addition, we will use the *interval Newton’s method*, a powerful extension of the iterative root-finding method which provides rigorous bounds on the zeros of a differentiable function and gives a sufficient condition to guarantee the function has a unique zero in a given interval [Tuc11, Ch. 5].

First note that  $f(0) = 0$  and  $f(x) > 0$  using the simple upper bound  $M(x) < \frac{1}{x}$  for  $x > 0$ . Hence, any maximum of  $f$  cannot occur at zero. Next,  $f$  is smooth, with first derivative

$$f'(x) = 1 + x^2 - x(2 + x^2)M(x)$$

and second derivative

$$f''(x) = x^3 + 4x - M(x)(2 + 5x^2 + x^4).$$

By using the interval Newton's method as implemented in the Julia programming language [BEKS17] package `IntervalRootFinding.jl` [BS19b], we can verify that for  $x \in [0.0, 3.0]$ , the equation  $f'(x) = 0$  has a unique solution  $x_0$  which satisfies (5.44). Moreover, bounding  $f$  on the interval given by (5.44) with interval arithmetic, as implemented in `IntervalArithmetic.jl` [BS19a] shows that  $\mu := f(x_0)$  satisfies (5.45). Lastly, we likewise find that

$$-0.16730889431005824 \leq f''(x_0) \leq -0.16730889430876594,$$

and hence  $f''(x_0) < 0$  confirming that  $x_0$  is a local maxima of  $f$ . The code used to establish these bounds can be found here: [Han20a]. This code uses the MPFR library [FHLP+07] for a correctly-rounded implementation of the complementary error function,  $1 - \Phi(x)$ .

Lastly, for  $x > 0$ , we use the lower bound  $M(x) > \frac{x}{x^2+1}$  which holds for  $x > 0$  [YC15]. This bound yields  $f(x) < \frac{x}{1+x^2}$ . The right-hand side is strictly monotone decreasing for  $x > 1$ , and evaluates to 0.3 at  $x = 3$ . Hence  $f(x) < 0.3$  for  $x > 3$ . Thus, the local maximum at  $x_0$  is in fact a global maximum.

## 5.6 Smoothed entropies and MaxEnt

Theorem 3.2.1 immediately yields maximizers and minimizers over  $B_\varepsilon(\rho)$  for any Schur concave function  $\varphi$ . Note that, as stated in the following corollary, the minimizer  $\rho_\varepsilon^*$  (resp. the maximizer  $\rho_{*,\varepsilon}$ ) in the majorization order (3.7) is the maximizer (resp. minimizer) of  $\varphi$  over the  $\varepsilon$ -ball  $B_\varepsilon(\rho)$ .

**Corollary 5.6.1.** *Let  $\varphi : \mathcal{P} \rightarrow \mathbb{R}$  be Schur concave. Then for  $p \in \mathcal{P}$ ,*

$$p_\varepsilon^* \in \operatorname{argmax}_{B_\varepsilon(p)} \varphi, \quad \text{and} \quad p_{*,\varepsilon} \in \operatorname{argmin}_{B_\varepsilon(p)} \varphi.$$

*Additionally, if  $\varphi$  is strictly Schur concave, any other state  $p' \in \operatorname{argmax}_{B_\varepsilon(p)} \varphi$  (resp.  $p' \in \operatorname{argmin}_{B_\varepsilon(p)} \varphi$ ) is unitarily equivalent to  $p_\varepsilon^*$  (resp.  $p_{*,\varepsilon}$ ). If  $\varphi$  is strictly concave, then  $\operatorname{argmax}_{B_\varepsilon(p)} \varphi = \{p_\varepsilon^*\}$ .*

*This result can be equivalently stated in terms of quantum states by substituting  $\rho \in \mathcal{D}(\mathcal{H})$  for  $p$  in the above if  $\varphi : \mathcal{D}(\mathcal{H}) \rightarrow \mathbb{R}$  is Schur concave.*

In an experimental setup, one may not know a quantum state exactly, and need to determine an estimate of the state for mathematical analysis. Let us briefly investigate Corollary 5.6.1's implications for this task. The so-called MaxEnt (or maximum-entropy) principle gives that an appropriate estimate of  $\sigma$  is one which is compatible with the constraints on  $\sigma$  and which has maximum entropy subject to those constraints [BDDA+99; Jay57].

Consider an experimental device which attempts to produce a given target pure state  $\sigma_{\text{target}}$ , and let  $\sigma$  denote the actual state produced by the device. One can estimate the fidelity between  $\sigma$  and  $\sigma_{\text{target}}$  efficiently, using few Pauli measurements of  $\sigma$  [FL11; dSLP11]. Using the bound  $\frac{1}{2}\|\sigma - \sigma_{\text{target}}\|_1 \leq \varepsilon := \sqrt{1 - F(\sigma, \sigma_{\text{target}})^2}$ , one therefore obtains a bound on the trace distance between  $\sigma$  and  $\sigma_{\text{target}}$ . Theorem 3.2.1 gives that the state with maximum entropy in  $B_\varepsilon(\sigma_{\text{target}})$  is  $\rho_\varepsilon^*(\sigma_{\text{target}})$ . Using that  $\sigma_{\text{target}}$  is a pure state, Theorem 4.2.2 yields

$$(\sigma_{\text{target}})_\varepsilon^* = \begin{cases} \text{diag}(1 - \varepsilon, \frac{\varepsilon}{d-1}, \dots, \frac{\varepsilon}{d-1}) & \text{if } \varepsilon \leq 1 - \frac{1}{d} \\ \tau := \frac{1}{d} & \text{else} \end{cases} \quad (5.51)$$

in the basis in which  $\sigma_{\text{target}} = \text{diag}(1, 0, \dots, 0)$ , where  $d$  is the dimension of the underlying Hilbert space. The maximum-entropy principle therefore implies that  $(\sigma_{\text{target}})_\varepsilon^*$  defined by (5.51) is the appropriate estimate of  $\sigma$ , when given only the condition that  $\frac{1}{2}\|\sigma - \sigma_{\text{target}}\|_1 \leq \varepsilon$ .

*Remarks.*

- it may be possible to determine additional constraints on  $\sigma$  by the Pauli measurements performed to estimate  $F(\sigma, \sigma_{\text{target}})$ . In that case, MaxEnt gives that the appropriate estimate of  $\sigma$  is the state with maximum entropy subject to these constraints as well, and not only the relation  $\frac{1}{2}\|\sigma - \sigma_{\text{target}}\|_1 \leq \varepsilon$ .
- it may be possible to devise a measurement scheme to estimate  $\frac{1}{2}\|\rho - \sigma_{\text{target}}\|_1$  directly, which could be more efficient than first estimating the fidelity and then employing the Fuchs-van de Graaf inequality to bound the trace distance.

Moreover, Corollary 5.6.1 yields maximizers and minimizers of any Schur concave function, not simply the von Neumann entropy. This allows computation of (trace-ball) “smoothed” Schur-concave functions. Given  $\varphi : \mathcal{D} \rightarrow \mathbb{R}$ , we define

$$\bar{\varphi}^{(\varepsilon)}(\sigma) := \max_{\omega \in B_\varepsilon(\rho)} \varphi(\omega), \quad \underline{\varphi}^{(\varepsilon)}(\sigma) := \min_{\omega \in B_\varepsilon(\rho)} \varphi(\omega). \quad (5.52)$$

By Corollary 5.6.1, we therefore obtain explicit formulas:  $\bar{\varphi}^{(\epsilon)}(\sigma) = \varphi(\rho_\epsilon^*)$ , and  $\underline{\varphi}^{(\epsilon)}(\sigma) = \varphi(\rho_{*,\epsilon})$ . In particular, this provides an exact version of Theorem 1 of [Sko16], which formulates approximate maximizers for the smoothed  $\alpha$ -Rényi entropy  $\bar{S}_\alpha^{(\epsilon)}$ .

Note that by setting  $\varphi = S_{\min}$  or  $S_{\max}$ , the min- and max-entropies, in (5.52), yields explicit expressions for the min- and max- entropies smoothed over the  $\epsilon$ -ball. These choices are of particular interest, due to their relevance in one-shot information theory (see e.g. [Ren05; Tom16] and references therein). In particular, let us briefly consider one-shot classical data compression. Given a source (random variable)  $X$ , one wishes to encode output from  $X$  using codewords of a fixed length  $\log m$ , such that the original message may be recovered with probability of error at most  $\epsilon$ . It is known that the minimal value of  $m$  at fixed  $\epsilon$ , denoted  $m_*(\epsilon)$ , satisfies

$$\underline{S}_{\max}^{(\epsilon)}(X) \leq \log m_*(\epsilon) \leq \inf_{\delta \in (0, \epsilon)} [\underline{S}_{\max}^{(\delta)}(X) + \log \frac{1}{\delta}] \quad (5.53)$$

as shown by [Tom16; RR12]. Equation (5.52) provides the means to explicitly evaluate the quantity  $\underline{S}_{\max}^{(\epsilon)}(X)$  in this bound.

*Remark.* Although Part I of this thesis is concerned with the  $\epsilon$ -ball defined via the trace distance, these results have some implications for other distance measures. For example, the so-called *sine distance* [Ras02; Ras06] or *purified distance* [Tom12] is often used in defining smoothed entropies. The purified distance between two states  $\rho$  and  $\sigma$ , denoted  $P(\rho, \sigma)$ , satisfies the inequality

$$T(\rho, \sigma) \leq P(\rho, \sigma) \quad (5.54)$$

by Lemma 3.5 of [Tom16]. Therefore, denoting  $\tilde{B}_\epsilon(\rho)$  by the  $\epsilon$ -ball

$$\tilde{B}_\epsilon(\rho) := \{\omega \in \mathcal{D}(\mathcal{H}) : P(\omega, \rho) \leq \epsilon\} \quad (5.55)$$

one has  $\tilde{B}_\epsilon(\rho) \subseteq B_\epsilon(\rho)$ . Therefore, for any Schur concave function  $f$ ,

$$f(\rho_{*,\epsilon}) \leq \min_{\tilde{B}_\epsilon(\rho)} f \leq \max_{\tilde{B}_\epsilon(\rho)} f \leq f(\rho_\epsilon^*) \quad (5.56)$$

by Corollary 5.6.1. We therefore obtain immediate bounds on the maxima and minima of Schur concave functions over  $\tilde{B}_\epsilon(\rho)$ .

## 5.7 Application to LOCC transformations

Nielsen's theorem [Nie99] states that one bipartite pure state  $\psi_{AB} \in \mathcal{D}_{\text{pure}}(\mathcal{H}_A \otimes \mathcal{H}_B)$  may be converted into another  $\phi_{AB} \in \mathcal{D}_{\text{pure}}(\mathcal{H}_A \otimes \mathcal{H}_B)$  by a protocol consisting of *local operations and classical communication* (LOCC) if and only if the reduced states  $\rho_A := \text{tr}_B[\psi_{AB}]$  and  $\sigma_A := \text{tr}_B[\phi_{AB}]$  satisfy  $\rho_A \prec \sigma_A$ . We write  $\psi_{AB} \rightarrow \phi_{AB}$  to denote that  $\psi_{AB}$  can be transformed into  $\phi_{AB}$  by an LOCC protocol, and call  $\psi_{AB}$  the *source state* and  $\phi_{AB}$  the *target state*. For the sake of brevity, we will omit the full mathematical definition of an LOCC protocol in this text; see e.g. [CLMO+14] as a reference.

In the case that  $\rho_A \not\prec \sigma_A$ , can an LOCC protocol perform a transformation of the source state to an approximation of the target state,  $\psi_{AB} \rightarrow \omega_{AB} \approx \phi_{AB}$ ? The question, in fact, has an optimal solution with respect to two different distance measures.

[VJN00] construct an optimal state  $\chi_{AB}^F \in \mathcal{D}_{\text{pure}}(\mathcal{H}_A \otimes \mathcal{H}_B)$  such that

$$\chi_{AB}^F \in \underset{\substack{\chi_{AB} \in \mathcal{D}_{\text{pure}}(\mathcal{H}_A \otimes \mathcal{H}_B) \\ \psi_{AB} \rightarrow \chi_{AB}}}{\text{argmax}} F(\chi_{AB}, \phi_{AB})^2$$

where  $F(\rho, \sigma) := \|\sqrt{\rho}\sqrt{\sigma}\|_1$  is the *fidelity* of two states  $\rho$  and  $\sigma$ . In the case of pure states  $\chi$  and  $\phi$ , the fidelity can be expressed as  $F(\chi, \phi) = |\langle \chi | \phi \rangle|$ . The fidelity can be related to the trace distance via the Fuchs and van de Graaf inequalities,

$$1 - F(\rho, \sigma) \leq T(\rho, \sigma) \leq \sqrt{1 - F(\rho, \sigma)^2}. \quad (5.57)$$

In the case that  $\rho$  and  $\sigma$  are both pure states, the second inequality is an equality. Thus, maximizing the fidelity between pure states is equivalent to minimizing the trace distance, and the state  $\chi_{AB}^F$  is likewise optimal with respect to the trace distance.

[CGV13] define a metric  $d$  on  $\mathcal{P}_d$  (where  $d := \dim \mathcal{H}_A$ ) by

$$d(p, q) = H(p) + H(q) - 2H(\sup_{\prec}(p, q))$$

where  $\sup_{\prec}(p, q)$  is the majorization-supremum discussed in Section 3.4.3. Let  $\vec{\lambda}(\rho) := (\lambda_i^\downarrow(\rho))_i \in \mathcal{P}_d^\downarrow$  be the sorted eigenvalues of  $\rho_A = \text{tr}_B[\psi_{AB}]$  and likewise  $\vec{\lambda}(\sigma) := (\lambda_i^\downarrow(\sigma))_i \in \mathcal{P}_d^\downarrow$  be the sorted eigenvalues of  $\sigma_A = \text{tr}_B[\phi_{AB}]$ . Define  $x = \sup_{\prec}(\vec{\lambda}(\rho), \vec{\lambda}(\sigma))$ , and let  $\phi_{AB}$  have the Schmidt decomposition

$$|\phi\rangle_{AB} = \sum_i \sqrt{\lambda_i^\downarrow(\sigma)} |i\rangle_A |i\rangle_B \quad (5.58)$$

for some basis  $\{|i\rangle_A\}_i$  of  $\mathcal{H}_A$  and  $\{|i\rangle_B\}_i$  of  $\mathcal{H}_B$ . Then [BSFH+17] shows that  $\chi_{AB}^{\sup} \in \mathcal{D}_{\text{pure}}(\mathcal{H}_A \otimes \mathcal{H}_B)$  given by

$$|\chi^{\sup}\rangle_{AB} = \sum_i \sqrt{x_i} |i\rangle_A \otimes |i\rangle_B \quad (5.59)$$

satisfies

$$\chi_{AB}^{\sup} \in \underset{\substack{\chi_{AB} \in \mathcal{D}_{\text{pure}}(\mathcal{H}_A \otimes \mathcal{H}_B) \\ \psi_{AB} \rightarrow \chi_{AB}}}{\text{argmin}} \quad d(\chi_{AB}, \phi_{AB}).$$

In other words, taking the majorization-supremum of the Schmidt coefficients gives the closest target state to  $\phi_{AB}$  according to the metric  $d$  that can be reached from  $\psi_{AB}$  by an LOCC protocol.

Here, we provide a modest contribution along the same lines, although without establishing optimality. [HOS18] showed that for  $p, q \in \mathcal{P}_d$  with  $p \not\prec q$ , the quantity

$$\delta^*(p, q) := 2 \max_{k \in \{1, \dots, d\}} \sum_{i=1}^k (q_i^\downarrow - p_i^\downarrow)$$

satisfies

$$\delta^*(p, q) = \min\{\delta_1 \geq 0 : q \prec p_{*, \delta_1}\} = \min\{\delta_2 \geq 0 : q_{\delta_2}^* \prec p\} \quad (5.60)$$

where  $p_{*, \delta_1}$  is the majorization-maximizer of  $B_{\delta_1}(p)$  and  $q_{\delta_2}^*$  is the majorization-minimizer of  $B_{\delta_2}(q)$ , as defined in Chapter 3.

Then, similarly to the construction (5.59), define

$$|\chi_*\rangle_{AB} := \sum_i \sqrt{(q_{\delta, *})_i} |i\rangle_A |i\rangle_B \quad (5.61)$$

where  $\{|i\rangle_A\}_i$  and  $\{|i\rangle_B\}_i$  are defined by the Schmidt decomposition of  $\phi_{AB}$  in (5.58), and where  $q := \vec{\lambda}(\sigma)$ , and  $\delta := \delta^*(\vec{\lambda}(\rho), \vec{\lambda}(\sigma))$ . Likewise, define

$$|\chi^*\rangle_{AB} := \sum_i \sqrt{(p_\delta^*)_i} |f_i\rangle_A |f_i\rangle_B \quad (5.62)$$

where  $\{|f_i\rangle_A\}_i$  and  $\{|f_i\rangle_B\}_i$  are orthonormal bases defined by the Schmidt decomposition of  $\psi_{AB}$ , and  $p = \vec{\lambda}(\rho)$ . Then, since  $p \prec q_{\delta, *}$ , we have  $\psi_{AB} \rightarrow (\chi_*)_{AB}$  and since  $p_\delta^* \prec q$ , we have  $(\chi^*)_{AB} \rightarrow \phi_{AB}$ .

Thus,  $(\chi_*)_{AB}$  serves as an approximation to the target state such that the LOCC transformation is possible, while likewise  $(\chi^*)_{AB}$  serves as an approximation to the source

state. We can quantify this approximation by bounding the trace distance between  $(\chi^*)_{AB}$  and  $\psi_{AB}$ .

**Proposition 5.7.1.** *Let  $\phi_{AB} \in \mathcal{D}_{\text{pure}}(\mathcal{H}_A \otimes \mathcal{H}_B)$  and  $\psi_{AB} \in \mathcal{D}_{\text{pure}}(\mathcal{H}_A \otimes \mathcal{H}_B)$  be pure states with reduced states given by  $\rho_A := \text{tr}_B[\psi_{AB}]$  and  $\sigma_A := \text{tr}_B[\phi_{AB}]$ . The state  $(\chi^*)_{AB}$  defined in (5.62) satisfies  $(\chi^*)_{AB} \rightarrow \phi_{AB}$  and the bound*

$$T((\chi^*)_{AB}, \psi_{AB}) \leq \sqrt{2\delta}$$

while  $(\chi_*)_{AB}$  defined in (5.61) satisfies  $\psi_{AB} \rightarrow (\chi_*)_{AB}$  and the bound

$$T((\chi_*)_{AB}, \phi_{AB}) \leq \sqrt{2\delta},$$

where  $\delta := \delta^*(\vec{\lambda}(\rho), \vec{\lambda}(\sigma))$ , which is defined in (5.60).

*Remark.* The state  $(\chi_*)_{AB}$  is not the optimal approximation to  $\phi_{AB}$  in trace distance that can be obtained from  $\psi_{AB}$  by an LOCC protocol; that state is  $(\chi^F)_{AB}$  constructed by [VJN00], as discussed above. By using the majorization-minimizer and maximizer, however, one can approximate either the source state and target state, as shown above.

*Proof.* We will establish the bound on  $T((\chi^*)_{AB}, \psi_{AB})$ ; the bound on  $T((\chi_*)_{AB}, \phi_{AB})$  follows in the same manner. We have that the fidelity between  $(\chi^*)_{AB}$  and  $\psi_{AB}$  satisfies

$$F((\chi^*)_{AB}, \psi_{AB}) = |\langle \chi^* | \psi \rangle| = \sum_i \sqrt{(p_\delta^*)_i p_i}.$$

This is precisely the fidelity between the reduced states

$$\rho_\delta^* = \text{diag}(p_\delta^*), \quad \rho = \text{diag}(p).$$

The Fuchs and van de Graaf inequalities, (5.57), give

$$F(\rho_\delta^*, \rho) \geq 1 - T(\rho_\delta^*, \rho) \tag{5.63}$$

Since  $(\chi^*)_{AB}$  and  $\psi_{AB}$  are pure states, they achieve equality in the upper bound of Fuchs and van de Graaf, namely  $T((\chi^*)_{AB}, \psi_{AB}) = \sqrt{1 - F((\chi^*)_{AB}, \psi_{AB})^2}$ . Thus,

$$\begin{aligned} T((\chi^*)_{AB}, \psi_{AB}) &= \sqrt{1 - F((\chi^*)_{AB}, \psi_{AB})^2} \\ &= \sqrt{1 - F(\rho_\delta^*, \rho)^2} \\ &\leq \sqrt{1 - (1 - T(\rho_\delta^*, \rho))^2} \end{aligned}$$



$$\leq \sqrt{2T(\rho_\delta^*, \rho)}.$$

using (5.63) for the first inequality. The proof is completed by the observation that  $T(\rho_\delta^*, \rho) \leq \delta$  since  $\rho_\delta^* \in B_\delta(\rho)$ .  $\square$

## 5.8 A continuity bound for $\rho \mapsto D(\rho\|\sigma)$

The Audenaert-Fannes bound (1.1) can be interpreted as a uniform continuity bound for  $\rho \mapsto D(\rho\|\frac{1}{d}\mathbf{1})$ , using that  $S(\rho) = \log d - D(\rho\|\frac{1}{d}\mathbf{1})$ , where  $D(\rho\|\sigma) = \text{tr}[\rho \log \rho - \rho \log \sigma]$  is the quantum relative entropy between two quantum states  $\rho, \sigma \in \mathcal{D}(\mathcal{H})$  with  $d := \dim \mathcal{H}$ .

We can easily generalize this bound to allow arbitrary full-rank states  $\omega$  in the second slot, instead of the maximally mixed state  $\frac{1}{d}\mathbf{1}$ , as follows. Note that this is not an application of majorization flow.

**Proposition 5.8.1.** *For any  $\rho, \sigma \in \mathcal{D}(\mathcal{H})$  and  $\omega > 0$ , we have*

$$|D(\rho\|\omega) - D(\sigma\|\omega)| \leq |S(\rho) - S(\sigma)| + T(\rho, \sigma)(\log \lambda_+(\omega) - \log \lambda_-(\omega)). \quad (5.64)$$

*Applying the Audenaert-Fannes bound to the first term yields that for  $\varepsilon \in [0, 1 - \frac{1}{d}]$  and  $\omega > 0$  the bound*

$$\sup_{\substack{\rho, \sigma \in \mathcal{D}(\mathcal{H}) \\ T(\rho, \sigma) \leq \varepsilon}} |D(\rho\|\omega) - D(\sigma\|\omega)| \leq \varepsilon \log(d-1) + h(\varepsilon) + \varepsilon(\log \lambda_+(\omega) - \log \lambda_-(\omega)) \quad (5.65)$$

*holds. Moreover, this bound is tight in the sense that for any  $\varepsilon \in [0, 1 - \frac{1}{d}]$  and any numbers  $\lambda_\pm > 0$  (with  $\lambda_+ \geq \lambda_-$ ) there exists a positive operator  $\omega$  with smallest and largest eigenvalues given by  $\lambda_\pm(\omega) \equiv \lambda_\pm$  and such that equality holds in (5.65).*

*Proof.* We have that

$$D(\rho\|\omega) - D(\sigma\|\omega) = \text{tr}(\rho \log(\rho)) - \text{tr}(\rho \log \omega) - \text{tr}(\sigma \log \sigma) + \text{tr}(\sigma \log \omega) \quad (5.66)$$

$$= S(\sigma) - S(\rho) + \text{tr}((\sigma - \rho) \log \omega). \quad (5.67)$$

Let  $(\sigma - \rho)_\pm$  denote the positive and negative parts of the operator  $\sigma - \rho$  (sometimes called the Jordan decomposition). Then

$$D(\rho\|\omega) - D(\sigma\|\omega) = S(\sigma) - S(\rho) + \text{tr}((\sigma - \rho)_+ \log \omega) - \text{tr}((\sigma - \rho)_- \log \omega). \quad (5.68)$$

Since  $\lambda_-(\omega)\mathbf{1} \leq \omega \leq \lambda_+(\omega)\mathbf{1}$ , and  $(\sigma - \rho)_\pm \geq 0$  in semi-definite order, we have

$$D(\rho\|\omega) - D(\sigma\|\omega) \leq S(\sigma) - S(\rho) + \text{tr}((\sigma - \rho)_+) \log \lambda_+(\omega) - \text{tr}((\sigma - \rho)_-) \log \lambda_-(\omega). \quad (5.69)$$

Since  $\text{tr}(\sigma - \rho) = 0$ , we have  $\text{tr}((\sigma - \rho)_\pm) = T(\rho, \sigma)$ , and (5.64) follows.

Next, choose  $\rho = \text{diag}(1, 0, \dots, 0)$ , a pure state which is diagonal in the computational basis, and let  $\sigma = \text{diag}(1 - \varepsilon, \frac{\varepsilon}{d-1}, \dots, \frac{\varepsilon}{d-1})$ . We have that  $T(\rho, \sigma) = \varepsilon$  and that  $S(\sigma) - S(\rho) = \varepsilon \log d + h(\varepsilon)$ . In fact, any two states saturating the Audenaert-Fannes inequality are of this form. Next, let  $\omega = \text{diag}(\lambda_-(\omega), \lambda_+(\omega), \dots, \lambda_+(\omega))$ . Then  $(\sigma - \rho)_+ = \text{diag}(0, \frac{\varepsilon}{d-1}, \dots, \frac{\varepsilon}{d-1})$  while  $(\sigma - \rho)_- = \text{diag}(\varepsilon, 0, \dots, 0)$ , so that

$$(\sigma - \rho)_\pm \log \omega = (\sigma - \rho)_\pm \log(\lambda_\pm(\omega)).$$

Thus, equality holds in Equation (5.69) for these states. This completes the proof.  $\square$

*Remark.* We have not assumed the normalization  $\text{tr}(\omega) = 1$ ; while this generalizes the inequality, the tightness proof only goes through in the case  $\lambda_- + (d-1)\lambda_+ = 1$ . Note also that in the case  $\omega = \mathbf{1}$ , we recover the Audenaert-Fannes bound.

# Chapter 6

## Guesswork

In this chapter, we consider a somewhat different entropic quantity: the guesswork. In Section 6.1 we review an interpretation of the Shannon entropy as the expected number of guesses required to win a certain binary guessing game. In Section 6.2, we define the guesswork as the expected number of guesses required to win a modified version of the guessing game, and establish a tight Lipschitz continuity bound for the quantity. In Section 6.3, we discuss a variant of the guesswork in which the player has access to quantum side information. In that section, we discuss the equivalence of various guessing strategies, provide several formulations of the quantity as optimization problems, establish its concavity and Lipschitz continuity, and establish some entropic bounds.

### 6.1 The Shannon entropy in a binary guessing game

In this section, we review an interpretation of the Shannon entropy in terms of a guessing game<sup>1</sup>. Consider a random variable  $X \sim p_X$  on some finite alphabet  $\mathcal{X}$ . The game has two participants, a referee and a player. The referee, Alice, draws a sample from the random variable and finds that  $X = x^*$  for some  $x^* \in \mathcal{X}$ . The player, Bob, knows the distribution  $p_X$  but not  $x^*$ , and is allowed to ask the referee binary questions of the form “Is  $x^* \in \tilde{\mathcal{X}}$ ?” for any subset  $\tilde{\mathcal{X}} \subseteq \mathcal{X}$ . This repeats until Bob asks such a question with  $|\tilde{\mathcal{X}}| = 1$  and the answer is ‘YES’, in which case he’s learned the true outcome.

We wish to understand: what is the minimal expected number of guesses required? As we will see, this quantity is characterized in some sense by the Shannon entropy. First, we may assume Bob’s strategy is deterministic: Bob starts with the same initial question, and given the answer, always chooses his second question according to some preformulated

---

<sup>1</sup>This interpretation is briefly described in [LHL07] and also appears in [Mac03, Chapter 4.1]

strategy, etc. Otherwise, if Bob's strategy is random, we can see it as drawing a random variable with outcomes in (the finitely many) possible deterministic strategies, and then following that strategy. Since the expectation is more than the minimum, he might as well choose an optimal deterministic strategy to start with. We also assume that if Bob learns  $x^* \notin \tilde{\mathcal{X}}$ , he will not ask if  $x^*$  is in some subset of  $\tilde{\mathcal{X}}$  again.

We can see Bob's *guessing strategy*, i.e. his sequence of questions, as conducting a binary search algorithm; if he chooses a different sequence of questions (a different guessing strategy), then he conducts this binary search in a different manner. We can see this very explicitly by organizing his questions into a binary tree. The root of the tree corresponds to the whole alphabet  $\mathcal{X}$ . Then Bob's first question corresponds to a partition of  $\mathcal{X}$  into some set  $\tilde{\mathcal{X}}_1$  and  $\mathcal{X} \setminus \tilde{\mathcal{X}}_1$ , which are the two nodes on the second level, and so forth. The leaves of the tree are the singleton elements  $\{\{x\} : x \in \mathcal{X}\}$ . Using this binary tree, we may encode a particular sequence of answers as a binary string, where 'YES' answers correspond to '0' (and moving to the left child in the tree), while 'NO' answer correspond to '1' (and moving to the right child in the tree). Such a sequence of answers terminates when a leaf is reached in the tree; equivalently, when Bob knows  $x^*$  lies in a singleton set and can then make the correct singleton guess on the next question.

We can interpret this in the language of *source coding*<sup>2</sup>. A (binary) code  $C$  for  $X$  is a map from  $\mathcal{X} \rightarrow \{0, 1\}^*$ , the set of finite binary strings. The map (the *code*) in this case is determined by Bob's sequence of guesses as above. The set of *codewords* is the image  $C(\mathcal{X})$ . The length of a codeword, namely the length of the binary string, corresponds exactly to the number of questions Bob had to ask in order to determine  $x^*$ . For  $x \in \mathcal{X}$ , we denote  $l(x)$  as the length of  $C(x)$ . We are interested, therefore, in the expected length  $L = \mathbb{E}[l(X)]$  of the codewords, with respect to the randomness in drawing  $x^*$  from  $X$ . See Figure 6.1 for an example guessing strategy and some of the associated lengths.

The code  $C$  described above has the special property that it is a *prefix-free code*, meaning no codeword is a prefix of another codeword. To see this, note that every codeword describes a path from the root of the binary tree to a leaf; a prefix of length  $m$  of a codeword corresponds to descending  $m$  levels of the tree by choosing to travel to either the left or right child at each step. But each codeword ends in a leaf, and there are no more children to travel to; hence, one codeword cannot be a prefix of another.

The Kraft inequality is a fundamental result about prefix-free codes which states that

$$\sum_{x \in \mathcal{X}} 2^{-l(x)} \leq 1.$$

<sup>2</sup>See e.g. [CT06, Chapter 5] for a reference for the source coding results quoted in this section.

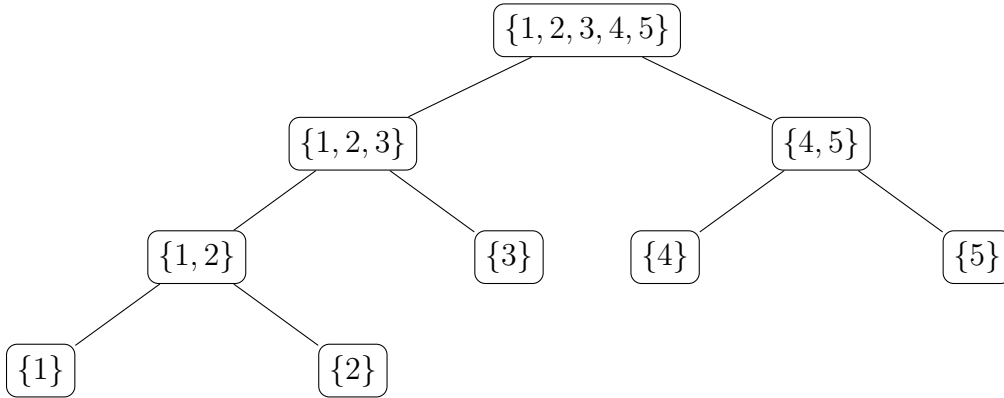


Fig. 6.1 An example guessing strategy with  $\mathcal{X} = \{1, 2, 3, 4, 5\}$ . With his first guess, Bob asks whether or not  $x^* \in \{1, 2, 3\}$ . Let's assume the answer is 'YES'; then his second guess asks whether or not  $x^* \in \{1, 2\}$ . Let's assume the answer is 'NO'. Then  $x^* = 3$ , and the sequence of guesses is encoded as 01. In this case, it only took two guesses to correctly guess  $x^* = 3$ ; we can see that guessing 4 or 5 will similarly be accomplished in two guesses. With this strategy, if  $x^* = 1$  or  $x^* = 2$ , however, it will take three guesses to determine them. Hence, this strategy will be better suited for a distribution with more weight on 3, 4, and 5, and less weight on 1 and 2.

The converse also holds: for any set of lengths satisfying the above inequality, there is a prefix-free code whose codewords are given by those lengths.

From this inequality, it follows that the expected length of the codewords satisfies

$$L \geq S(X)$$

with equality if and only if  $2^{-l(x)} = p_X(x)$ , and where  $S(X) = -\sum_{x \in \mathcal{X}} p_X(x) \log p_X(x)$  is the Shannon entropy (using base-2 logarithm).

A probability distribution  $p_X$  that can be written as  $p_X(x) = 2^{-l(x)}$  for some natural numbers  $\{l(x) : x \in \mathcal{X}\}$  is called 2-adic. For such a distribution, we may construct a prefix-free code with lengths given by  $l(x)$  (due to the converse of the Kraft inequality), and hence obtain a code with  $L = S(X)$ . In this sense, the Shannon entropy characterizes the minimum expected number of binary guesses needed to guess a realization of  $X$ .

This may not be entirely satisfactory, because for distributions that are not 2-adic, however, the inequality can never be saturated. However, by simply choosing  $l(x) = \lceil \log_2 \left( \frac{1}{p_X(x)} \right) \rceil$ , the Kraft inequality holds, and hence one can choose a prefix-free code with such lengths. Then its expected length satisfies

$$S(X) \leq L = \sum_{x \in \mathcal{X}} p_X(x) l(x) \leq S(X) + 1.$$

Using this, one can construct a code for  $N$  i.i.d. copies of  $X$  in which the expected length of the codebook,  $L_N$ , satisfies

$$S(X) \leq \frac{1}{N} L_N \leq S(X) + \frac{1}{N}$$

and in particular,  $\lim_{N \rightarrow \infty} \frac{1}{N} L_N = S(X)$ . So for general (non-2-adic) distributions, we at least recover the exact characterization of the Shannon entropy as the minimum expected number of binary guesses in this asymptotic sense.

*Remark.* Given a code for the source  $X$  with lengths  $\{l(x) : x \in \mathcal{X}\}$  which saturates the Kraft inequality, we can define a probability distribution  $q_X$  by  $q_X(x) = 2^{-l(x)}$  for  $x \in \mathcal{X}$ . Then if this distribution is used to encode symbols from a source with distribution  $p_X$ , we find

$$L = \mathbb{E}_p[l(X)] = S(p) + D(p||q)$$

where  $S(p) = -\sum_i p_i \log p_i$  is again the Shannon entropy, now regarded as a function of the distribution  $p$ . This provides an interpretation of the relative entropy  $D(p||q) := \sum_{i=1}^d p_i (\log p_i - \log q_i)$  as the number of additional bits needed to encode  $X$  using a code whose lengths are associated to distribution  $q$  instead of  $p$ .

## 6.2 The guesswork: A different guessing game

The guessing game discussed above provides one sense in which the Shannon entropy captures a notion of uncertainty about the random variable  $X$  on some finite alphabet  $\mathcal{X}$ . Intuitively, one is more uncertain about the value of  $X$  if it takes more guesses to figure out what that value is.

The binary guessing game described in the previous section, however, isn't the only way to quantify uncertainty through guessing. Let's change the rules. Instead of asking "Is  $x^* \in \tilde{\mathcal{X}}$ ?", Bob now may only ask questions of the form "Is  $x^* = x$ ?". The minimal expected number of guesses in this case is called the *guesswork* of  $X$ , denoted  $G(X)$ . This clearly provides another measure of the uncertainty in the value of  $X$ . It has a simple closed-form expression as

$$G(X) = \sum_{i=1}^{|\mathcal{X}|} i p_i^\downarrow \tag{6.1}$$

where  $X \sim p$ . In the following, we will denote  $G(X)$  equivalently in terms of the probability mass function  $p$  of  $X$  as  $G(p)$ .

Consider the familiar probability vector given by

$$p_\varepsilon^* = \left(1 - \varepsilon, \frac{\varepsilon}{|\mathcal{X}| - 1}, \dots, \frac{\varepsilon}{|\mathcal{X}| - 1}\right)$$

for some  $\varepsilon \in (0, 1 - \frac{1}{|\mathcal{X}|})$ , where  $p = (1, 0, \dots, 0)$ . Then the Shannon entropy satisfies

$$S(p_\varepsilon^*) = \varepsilon \log(|\mathcal{X}| - 1) + h_2(\varepsilon)$$

where  $h_2(\varepsilon) = -\varepsilon \log \varepsilon - (1 - \varepsilon) \log(1 - \varepsilon)$  is the binary entropy. The guesswork is given by

$$G(p_\varepsilon^*) = (1 - \varepsilon) \cdot 1 + \sum_{i=2}^{|\mathcal{X}|} \frac{i \cdot \varepsilon}{|\mathcal{X}| - 1} = 1 + \varepsilon \frac{|\mathcal{X}|}{2}.$$

We can see the guesswork scales linearly with  $|\mathcal{X}|$ , which makes sense because one essentially needs to perform a linear search over the tail of the distribution, while the entropy scales logarithmically with  $|\mathcal{X}|$ .

One might notice that in the case of the guesswork, the  $|\mathcal{X}|$ th guess, if it occurs, is unnecessary: after eliminating  $|\mathcal{X}| - 1$  outcomes, there can only be one possibility left. This motivates the so-called *modified guesswork*  $\tilde{G}$ , given by

$$\tilde{G}(p) := \sum_{i=1}^{|\mathcal{X}|-1} i \cdot p_i^\downarrow + (|\mathcal{X}| - 1)p_{|\mathcal{X}|}^\downarrow \quad (6.2)$$

[LHL07] showed that the modified guesswork is related to the Shannon entropy by a simple formula:

$$\tilde{G}(p) = H(p) + D(p^\downarrow \| \{2^{-c_i}\}_{i=1}^{|\mathcal{X}|}) \quad (6.3)$$

where  $D(p\|q) := \sum_{i=1}^{|\mathcal{X}|} p_i (\log p_i - \log q_i)$  is the relative entropy.

We may define a generalized guesswork  $G_{\vec{c}}(X)$  where  $\vec{c} = (c_i)_{i=1}^{|\mathcal{X}|} \in \mathbb{R}^{|\mathcal{X}|}$  is a *cost vector*, satisfying

$$0 \leq c_1 \leq c_2 \leq \dots \leq c_{|\mathcal{X}|}.$$

Then we define

$$G_{\vec{c}}(X) := \sum_{i=1}^{|\mathcal{X}|} c_i p_i^\downarrow. \quad (6.4)$$

Note that this definition unifies (6.1) and (6.2); choosing  $c_i = i$  for each  $i$  recovers the guesswork, while choosing

$$c_i = \begin{cases} i & i < d \\ |\mathcal{X}| - 1 & i = |\mathcal{X}| \end{cases}$$

recovers the modified guesswork. The generalized guesswork is a symmetric function of the distribution  $p$ , as it is only a function of the sorted entries of  $p$ . Moreover, we can write

$$G_{\vec{c}}(X) = \min_{\pi \in S_{|\mathcal{X}|}} \langle \pi(p), \vec{c} \rangle = \min_{\pi \in S_{|\mathcal{X}|}} \langle p, \pi^{-1}(\vec{c}) \rangle$$

As a minimum of linear functions,  $G_{\vec{c}}$  is a concave function of  $p$ . Since  $G_{\vec{c}}$  is symmetric and concave, it is Schur concave. Let us then investigate its continuity with the tools of Section 5.1. Note that  $G_{\vec{c}}$  is a generalization of the case considered in Example 4.3.5.

**Proposition 6.2.1.**  *$G_{\vec{c}}$  is Lipschitz continuous on  $\mathcal{P}$  with an optimal Lipschitz constant given by  $c_{|\mathcal{X}|} - c_1$ .*

*Proof.* While  $G_{\vec{c}}$  is not differentiable on  $\mathcal{P}_+$ , the function  $F = G_{\vec{c}}|_{\mathcal{P}^\downarrow}$  satisfies  $F(r) = \sum_{j=1}^{|\mathcal{X}|} c_j r_j$  and hence has an extension  $\tilde{F} : \mathcal{P} \rightarrow \mathbb{R}$  defined by the same formula. Since  $\tilde{F}$  is differentiable on  $\mathcal{P}_+$ , Corollary 4.3.6 shows that for  $p \in \mathcal{P}$ ,

$$G_{\vec{c}}(p_\varepsilon^*) - G_{\vec{c}}(p) = \int_0^\varepsilon \Gamma_{\tilde{F}}(\mathcal{M}_s(p^\downarrow)) \, ds.$$

Hence, as described in Section 4.4, bounding  $\Gamma_{\tilde{F}}(\mathcal{M}_s(p^\downarrow))$  provides a Lipschitz constant for  $G_{\vec{c}}$ . We have that

$$\Gamma_{\tilde{F}}(r^\downarrow) = \frac{1}{k_-} \sum_{i:r_i=r_-} c_i - \frac{1}{k_+} \sum_{i:r_i=r_+} c_i.$$

Since  $\frac{1}{k_-} \sum_{i:r_i=r_-} c_i \leq c_{|\mathcal{X}|}$  and  $\frac{1}{k_+} \sum_{i:r_i=r_+} c_i \geq c_1$ , we have that  $c_{|\mathcal{X}|} - c_1$  is a Lipschitz constant for  $G_{\vec{c}}$ . Moreover, if  $p \in \mathcal{P}^\downarrow$  has that  $p_\pm$  are both non-degenerate, and  $q = \mathcal{M}_\varepsilon(p)$  for  $\varepsilon < \delta(p)$  for  $\delta$  defined in Lemma 4.3.2, then that Lemma shows that  $q = (p_1 - \varepsilon, p_2, \dots, p_{|\mathcal{X}|-1}, p_{|\mathcal{X}|} + \varepsilon)$ , and we immediately have

$$G_{\vec{c}}(p_\varepsilon^*) - G_{\vec{c}}(p) = \varepsilon(c_{|\mathcal{X}|} - c_1).$$

Hence,  $c_{|\mathcal{X}|} - c_1$  is the optimal Lipschitz constant.  $\square$

*Remark.* Note that  $\Gamma_{\tilde{F}}$  defined in the above proof is not symmetric, and hence is not Schur convex on  $\mathcal{P}$ . Thus, this choice of  $\tilde{F}$  does not provide a path to establishing a uniform continuity bound by establishing the Schur convexity of  $\Delta_\varepsilon^{G_{\vec{c}}}$  as discussed in



Section 4.4. However, this does not rule out the Schur convexity of  $\Delta_\varepsilon^H$ . We may calculate for  $q = (1, 0, \dots, 0)$  and  $0 < \varepsilon < 1 - \frac{1}{|\mathcal{X}|}$ ,

$$\Delta_\varepsilon^{G_\varepsilon}(q) = G_\varepsilon(q_\varepsilon^*) - G_\varepsilon(q) = (1 - \varepsilon)c_1 + \frac{\varepsilon}{|\mathcal{X}| - 1} \sum_{i=2}^{|\mathcal{X}|} c_i - c_1 = \varepsilon \left( \frac{1}{|\mathcal{X}| - 1} \sum_{i=2}^{|\mathcal{X}|} c_i - c_1 \right).$$

On the other hand, by choosing  $p$  as in the proof of optimality of the Lipschitz constant for  $G_\varepsilon$ , we obtain  $\Delta_\varepsilon^{G_\varepsilon}(p) = \varepsilon(c_{|\mathcal{X}|} - c_1)$ . As long as for some  $i \in \{2, \dots, |\mathcal{X}| - 1\}$  one has  $c_i \neq c_{|\mathcal{X}|}$ , then  $\Delta_\varepsilon^{G_\varepsilon}(p) > \Delta_\varepsilon^{G_\varepsilon}(q)$ . Since  $p \prec q$ , we must have that  $\Delta_\varepsilon^{G_\varepsilon}$  is not Schur convex. Thus, in general, one cannot obtain a tight uniform continuity bound for  $G_\varepsilon$  via the Schur convexity of  $\Delta_\varepsilon^{G_\varepsilon}$ . However, since the bound

$$|G_\varepsilon(p) - G_\varepsilon(q)| \leq \text{TV}(p, q) \cdot (c_{|\mathcal{X}|} - c_1) \quad (6.5)$$

can in fact be achieved, the Lipschitz bound (6.5) itself provides a tight uniform continuity bound for  $G_\varepsilon$ .

Recall that in Section 5.1.1 it was shown that there exists a symmetric and concave function  $f$  such that its derivative along the path of majorization flow,  $\Gamma_f$ , is not Schur convex. The Schur convexity of  $\Gamma_H$  plays a role in the proof of continuity bounds for a quantity  $H$  in Section 4.4 by establishing the Schur convexity of its integrated form,  $\Delta_\varepsilon^H$ . Hence, since  $\Delta_\varepsilon^{G_\varepsilon}$  is not Schur convex in general, the guesswork provides a counterexample of a similar flavor: namely a quantity  $H$  being symmetric and concave is not enough to obtain the Schur convexity of  $\Delta_\varepsilon^H$ . In this case, however, the quantity of interest is not differentiable on  $\mathcal{P}_+$ .

## 6.3 The guesswork with quantum side information

Let us first briefly discuss the guesswork with classical side information. In this scenario, there are two random variables,  $X$  and  $Y$ . Bob knows the full joint distribution of  $XY$ . Alice obtains a realization  $X = x^*$ , and Bob learns the outcome of  $Y$ , namely  $Y = y$ . If the two random variables are correlated, Bob has learned something extra about  $X$ ; in fact, knowing  $Y = y$ , the distribution of  $X$  is given by  $p_{X|Y}(\cdot|y)$ . Then Bob can perform the optimal strategy to guess  $x^*$  (namely, sort this conditional distribution in descending order, and guess from most probable to least). The expected number of guesses required to obtain the correct answer is the *guesswork of  $X$  with classical side information  $Y$* ,

denoted  $G(X|Y)$ , and satisfies

$$G(X|Y) = \sum_{y \in \mathcal{Y}} p_Y(y) G(X|Y = y)$$

where  $\mathcal{Y}$  is the alphabet of  $Y$ , and  $G(X|Y = y) := G(p_{X|Y}(\cdot|y))$  is the guesswork of the conditional distribution. Like with the guesswork itself, this quantity can be generalized to use a cost vector  $\vec{c}$  by simply defining

$$G_{\vec{c}}(X|Y) = \sum_{y \in \mathcal{Y}} p_Y(y) G_{\vec{c}}(X|Y = y).$$

Now, consider the scenario in which when Alice obtains the realization  $X = x^*$ , Bob obtains a quantum state  $\rho_B^{x^*}$ . This occurs, for instance, if a quantum state is measured, and Alice has access to the measurement outcome while Bob has access to the post-measurement state. In this case, the joint distribution  $XY$  is replaced by a classical-quantum (c-q) state,

$$\rho_{XB} = \sum_{x \in \mathcal{X}} p_X(x) |x\rangle \langle x| \otimes \rho_B^x.$$

Given the state  $\rho_B^x$ , and knowing the joint c-q state  $\rho_{XB}$ , Bob aims to guess the index  $x \in \mathcal{X}$ . The expected number of guesses required to be successful, minimized over all guessing strategies, is the *guesswork with quantum side information*, denoted  $G(X|B)_\rho$ .

While the case of classical side information admits a very simple optimal strategy (which amounts to simply sorting the conditional probabilities  $p_{X|Y}(\cdot|y)$  in non-increasing order and guessing accordingly), the quantum case requires measurement on the quantum system  $B$ , which potentially disturbs the state of  $B$ , a priori complicating the analysis of the sequence of guesses in the optimal strategy. We show in Section 6.3.2, however, that a general sequential strategy is in fact equivalent to performing a single generalized measurement yielding a classical random variable  $Y$  of outcomes and then performing the optimal strategy using this  $Y$  as the classical side information. The earlier work [CCWF15] instead defined guesswork with quantum side information as the latter quantity, i.e., a measured version of the guesswork in the presence of classical side information. While these definitions are equivalent, we consider the definition in terms of a sequential protocol to be a more natural one. Moreover, the above-mentioned equivalence is proved via an explicit construction, allowing such a guessing strategy to be implemented sequentially. The single-measurement protocol could in general involve making a measurement with exponentially (in  $|\mathcal{X}|$ ) many outcomes. Hence it may be more efficient to implement it

instead as a sequence of (linearly-many) measurements with linearly-many outcomes, as allowed by the above construction.

This is related to the task of maximizing the “guessing probability”  $p_{\text{guess}}(X|B)$  [KRS09]. In that task, Bob is only given one attempt to guess the value of  $X$  and wishes to maximize the probability of doing so successfully. The guessing probability is related to the so-called *conditional min-entropy*  $H_{\min}(X|B)$  of the c-q state  $\rho_{XB}$ . In some sense, we can consider the guesswork to be an extension of the guessing probability. However, the nature of the optimization being done is different: instead of maximizing the probability of success in one attempt, we minimize the total number of guesses required. Therefore, the operations that a guesser performs to minimize the guesswork may be very different from those needed to maximize the guessing probability; for example, one might expect in the case in which only a single guess is being made, one may perform very “destructive” measurements to extract maximal information from the state at the cost of having a less informative post-measurement state, whereas in the scenario in which multiple measurements and multiple guesses are being made, the measurements may be chosen to be less destructive. Some of the connections between these two tasks have been investigated in [CCWF15].

We will see that the guesswork  $G(X|B)_\rho$  is concave in the state  $\rho_{XB}$ , but not unitarily invariant, just as for the conditional entropy  $H(X|B)_\rho$ . This means that the techniques of Section 5.1 cannot be applied to obtain a continuity bound for  $G(X|B)_\rho$ . Moreover, unlike for the guesswork (or the guesswork with classical side-information), we do not have a closed-form expression for  $G(X|B)_\rho$ , and the quantity is in general not differentiable (since one can see  $G(X)$  as a special case, and that quantity is not differentiable on all of  $\mathcal{P}_+$  as discussed in the previous section.) Thus, the techniques of Chapter 7 cannot be used to locally maximize  $G(X|B)_\rho$ , as was done in Corollary 7.3.2 for the conditional entropy, which otherwise might possibly provide a path towards establishing a (local) continuity bounds.

Instead, we will establish a semidefinite optimization problem (SDP) formulation of the guesswork with quantum side information. By studying the dependence of that formulation on the state  $\rho_{XB}$ , we quickly obtain a Lipschitz continuity bound for the quantity which has the feature of not depending on the dimension of the  $B$ -system (like the Alicki-Fannes bound [AF04], or Winter’s strengthening thereof [Win16], for the quantum conditional entropy).

**Overview of the remainder of the chapter** In Section 6.3.1, the guesswork task is described in more detail. In Section 6.3.2 a general framework of guessing strategies

is developed which allows formulating a guesswork task with arbitrary kinds of side information; additionally, several types of strategies with quantum side information are described and shown to be equivalent. In Section 6.3.3, the guesswork with quantum side information is described by a semidefinite optimization problem (SDP) representation. This is used to establish the concavity and Lipschitz continuity of the guesswork with quantum side information in Section 6.3.4. Lastly, in Section 6.3.5 techniques for computing the quantity are discussed, and in Section 6.3.6 entropic bounds are established.

### 6.3.1 Statement of the problem

Alice chooses a letter  $x \in \mathcal{X}$  with some probability  $p_X(x)$ , where  $\mathcal{X}$  is a finite alphabet. This naturally defines a random variable  $X \sim p_X(x)$ . She then sends a quantum system  $B$  to Bob, prepared in the state  $\rho_B^x$ , which depends on her choice  $x$ . Bob knows the set of states  $\{\rho_B^x : x \in \mathcal{X}\}$ , and the probability distribution  $\{p_X(x) : x \in \mathcal{X}\}$ , but he does not know which particular state is sent to him by Alice. Bob's task is to guess  $x$  correctly with as few guesses as possible. From Bob's perspective, he therefore has access to the  $B$ -part of the c-q state

$$\rho_{XB} = \sum_x p_X(x) |x\rangle\langle x|_X \otimes \rho_B^x. \quad (6.6)$$

In the purely classical case, this task reduces to the following scenario: Alice holds the random variable  $X \sim p_X(x)$ , and Bob holds a correlated random variable  $Y$ , and knows the joint distribution of  $(X, Y)$ . In this case  $\rho_{XB}$  reduces to the state

$$\rho_{XY} = \sum_x p_X(x) |x\rangle\langle x|_X \otimes \sum_y p_{Y|X}(y|x) |y\rangle\langle y|_Y. \quad (6.7)$$

In this case, if Bob's random variable  $Y$  has value  $y$ , then an optimal guessing strategy would be to sort the conditional distribution  $p_{X|Y}(\cdot|y)$  in non-increasing order so that

$$p_{X|Y}(x_1|y) \geq p_{X|Y}(x_2|y) \geq \dots \geq p_{X|Y}(x_{|\mathcal{X}|}|y) \quad (6.8)$$

and simply guess first  $x_1$ , then  $x_2$ , etc., until he gets it correct [Ari96].

In the case in which Bob's system  $B$  is quantum, he is allowed to perform any local operations he wishes on  $B$ , and then make a first guess  $x_1$ . He is told by Alice whether or not his guess is correct; then he can perform local operations on  $B$ , and make another guess, and so forth. We are interested in determining the minimal number of guesses needed on average for a given ensemble  $\{p_X(x), \rho_B^x\}_{x \in \mathcal{X}}$  and the associated optimal strategy.

More generally, we allow Bob to make  $K$  guesses, with possibly  $K < |\mathcal{X}|$ . Formally, we assume that Bob always makes all  $K$  guesses; any guess after the correct guess simply does not factor into the calculation of the minimal number of guesses (see Section 6.3.2 for a more detailed definition of the minimal number of guesses). Thus, Bob makes a sequence of guesses,  $g_1, \dots, g_K \in \mathcal{X}^K$  with some probability.

We could consider the scenario in which Bob makes a guess  $x_1$ , then learns whether or not the guess was correct, and uses that information to make his second guess  $x_2$ , and so forth. However, if Bob learns that his  $j$ th guess  $x_j$  is correct, then it does not matter what he guesses subsequently (it has no bearing on the minimal number of guesses). If the guess is incorrect, then his subsequent guesses do matter, and he should make his next guess accordingly. Hence, in such a protocol, the feedback about whether or not the  $j$ th guess is correct does not help, and Bob might as well assume that each guess is incorrect.

### 6.3.2 Guessing strategies

When Alice chooses  $x^* \in \mathcal{X}$ , a guessing strategy for Bob outputs a sequence of guesses  $\vec{g} = (g_1, \dots, g_K) \in \mathcal{X}^K$  with some probability  $p_{\vec{G}|X}(\vec{g}|x^*)$ . Hence, formally, a *guessing strategy for  $X$  with  $K$  guesses* is a random variable  $\vec{G}$  on  $\mathcal{X}^K$  that is correlated with  $X$ , such that  $(X, \vec{G})$  has marginal  $X \sim p_X$ . Note that the definition of a guessing strategy has no reference to the side information (if any) that Bob has access to; instead, the side information dictates the set of guessing strategies Bob has access to. This allows various types of side information to be analyzed within a uniform framework; in particular, the set of strategies available when Bob has access to some classical side information  $Y$  is described in Section 6.3.2, while the case of quantum side information is described in Section 6.3.2.

We are interested in the minimal number of guesses required to guess  $x^*$  correctly. This is defined as follows:

$$N(\vec{g}, x^*) := \begin{cases} \min \{j : g_j = x^*\} & g_j = x^* \text{ for some } j = 1, \dots, K \\ \infty & \text{else,} \end{cases} \quad (6.9)$$

where the outcome  $\infty$  occurs when none of the  $K$  guesses are correct. We can view  $N$  as a random variable taking values in  $\{1, 2, \dots, K, \infty\}$ . Given a guessing strategy  $\vec{G}$ , the quantity of interest is  $N(\vec{G}, X)$ , the corresponding random variable. We define

$$\mathcal{S}_K(X) := \left\{ N(\vec{G}, X) : X\vec{G} \sim p_{X\vec{G}} \right\} \quad (6.10)$$

to be the set of all possible random variables  $N$  associated to all guessing strategies  $\vec{G}$  with  $K$  guesses. We say two guessing strategies  $\vec{G}$  and  $\vec{G}'$  for  $X$  with  $K$  guesses are *equivalent* if  $N(\vec{G}', X) = N(\vec{G}, X)$ .

Note that if  $\vec{G}$  and  $\vec{G}'$  are two strategies with  $K$  guesses for  $X$  that differ only in guesses made after guessing the correct answer, then they are equivalent. This formalizes the notion introduced at the end of the previous section: since guesses made after the correct answer do not change the value of  $N(\vec{g}, x^*)$ , feedback of whether or not  $g_j = x^*$  can only lead to equivalent strategies.

### Classical strategies

Consider a pair of random variables  $(X, Y)$  where  $X$  has a finite alphabet  $\mathcal{X}$  and  $Y$  has a countable alphabet  $\mathcal{Y}$ . Alice chooses  $x^* \in \mathcal{X}$  (with probability  $p_X(x^*)$ ) and Bob is given  $y \in \mathcal{Y}$  (with probability  $p_{Y|X}(y|x^*)$ ). Bob's task is to guess  $x^*$ . Since Bob's sequence of guesses  $(g_1, \dots, g_K)$  can only depend on  $x^*$  via  $y$ , a classical guessing strategy  $\vec{G}$  is any random variable  $\vec{G}$  such that the ordered triple  $(X, Y, \vec{G})$  of random variables forms a Markov chain, which we denote as  $X - Y - \vec{G}$ . Hence, given a joint probability distribution  $p_{XY}$ , we define the set of random variables  $N$  associated to classical guessing strategies as follows:

$$\mathcal{S}_K^{\text{Classical}}(p_{XY}) := \left\{ N(\vec{G}, X) : X - Y - \vec{G} \right\} \subseteq \mathcal{S}_K(X). \quad (6.11)$$

### Equivalence of quantum strategies

Let us consider three classes of quantum strategies:

1. Measured strategy: Bob performs an arbitrary POVM  $\{E_y\}_{y \in \mathcal{Y}}$  on the  $B$ -system. Let  $Y$  be the random variable with outcomes in a finite alphabet  $\mathcal{Y}$  corresponding to his measurement outcomes, i.e.

$$p_{Y|X}(y|x) = \text{tr}[E_y \rho_B^x], \quad \forall x \in \mathcal{X}, y \in \mathcal{Y}. \quad (6.12)$$

Bob then employs a classical guessing strategy on  $(X, Y)$ . The set of random variables corresponding to the possible number of guesses under such a strategy is given by

$$\mathcal{S}_K^{\text{Measured}}(\rho_{XB}) := \left\{ N(\vec{G}, X) : X - Y - \vec{G}, Y \text{ satisfies (6.12)} \right. \\ \left. \text{for some finite alphabet } \mathcal{Y} \text{ \& POVM } \{E_y\}_{y \in \mathcal{Y}} \right\}. \quad (6.13)$$

We then observe that

$$\mathcal{S}_K^{\text{Measured}}(\rho_{XB}) \subseteq \mathcal{S}_K(X). \quad (6.14)$$

2. Ordered strategy: Bob performs a measurement with outcomes in  $\mathcal{X}^K$ , which are identified with guessing orders; i.e., if the outcome is  $(x_1, \dots, x_K) \in \mathcal{X}^K$ , Bob first guesses  $x_1$ , then  $x_2$ , and so forth. In this case, Bob performs a POVM  $\{E_{\vec{g}}\}_{\vec{g} \in \mathcal{X}^K}$  and the guessing strategy  $\vec{G}$  is distributed according to

$$p_{\vec{G}|X}(\vec{g}|x) = \text{tr}[E_{\vec{g}}\rho_B^x]. \quad (6.15)$$

As above, we define

$$\mathcal{S}_K^{\text{Ordered}}(\rho_{XB}) := \left\{ N(\vec{G}, X) : (\vec{G}, X) \text{ satisfy (6.15) for some POVM } \{E_{\vec{g}}\}_{\vec{g} \in \mathcal{X}^K} \right\} \quad (6.16)$$

It is evident that

$$\mathcal{S}_K^{\text{Ordered}}(\rho_{XB}) \subseteq \mathcal{S}_K^{\text{Measured}}(\rho_{XB}) \quad (6.17)$$

because any such ordered strategy is a special type of measured strategy (with  $Y = \vec{G}$ ). However, any measured strategy can in fact be simulated by an ordered strategy. Suppose we have a measured strategy with alphabet  $\mathcal{Y}$ , POVM  $\{E_y\}_{y \in \mathcal{Y}}$ , and  $\vec{G}$  satisfying  $X - Y - \vec{G}$ . Then

$$p_{\vec{G}|X}(\vec{g}|x) = \sum_{y \in \mathcal{Y}} p_{\vec{G}|Y}(\vec{g}|y) p_{Y|X}(y|x) = \sum_{y \in \mathcal{Y}} p_{\vec{G}|Y}(\vec{g}|y) \text{tr}[E_y \rho_B^x], \quad (6.18)$$

where we have used the Markov property for the first equality and (6.12) for the second equality.

Let  $\tilde{E}_{\vec{g}} := \sum_{y \in \mathcal{Y}} p_{\vec{G}|Y}(\vec{g}|y) E_y$ . Note  $\{\tilde{E}_{\vec{g}}\}_{\vec{g} \in \mathcal{X}^K}$  is a POVM: each element is positive semidefinite since  $\{E_y\}_{y \in \mathcal{Y}}$  is a POVM, and

$$\sum_{\vec{g} \in \mathcal{X}^K} \tilde{E}_{\vec{g}} = \sum_{\vec{g} \in \mathcal{X}^K} \sum_{y \in \mathcal{Y}} p_{\vec{G}|Y}(\vec{g}|y) E_y = \sum_{y \in \mathcal{Y}} \sum_{\vec{g} \in \mathcal{X}^K} p_{\vec{G}|Y}(\vec{g}|y) E_y = \sum_{y \in \mathcal{Y}} E_y = \mathbf{1}_B, \quad (6.19)$$

using again that  $\{E_y\}_{y \in \mathcal{Y}}$  is a POVM. Then substituting the definition of  $\tilde{E}_{\vec{g}}$  into (6.18) yields

$$p_{\vec{G}|X}(\vec{g}|x) = \text{tr}[\tilde{E}_{\vec{g}} \rho_B^x] \quad (6.20)$$

and hence (6.15) is satisfied with  $E = \tilde{E}$ . Therefore,

$$\mathcal{S}_K^{\text{Ordered}}(\rho_{XB}) = \mathcal{S}_K^{\text{Measured}}(\rho_{XB}). \quad (6.21)$$

3. Sequential quantum strategy: Suppose that Alice chooses  $x$  (which occurs with probability  $p_X(x)$ ), and hence Bob has the state  $\rho_B^x$ . To make his first guess, Bob chooses a set of generalized measurement operators  $\{M_x^{(1)}\}_{x \in \mathcal{X}}$  and reports the measurement outcome as his guess. He gets outcome  $x_1$  with probability

$$p_{G_1|X}(x_1|x) = \text{tr}[M_{x_1}^{(1)} \rho_B^x M_{x_1}^{(1)\dagger}] \quad (6.22)$$

and his post-measurement state is

$$\frac{1}{p_{G_1|X}(x_1|x)} M_{x_1}^{(1)} \rho_B^x M_{x_1}^{(1)\dagger}. \quad (6.23)$$

*Note:* in general, Bob could perform a unitary operation  $U_1$  on his state before measuring it. However, this would simply correspond to measuring with  $\{M_x^{(1)} U_1\}_{x \in \mathcal{X}}$  instead. Hence, it suffices to simply consider a generalized measurement  $\{M_x^{(1)}\}_{x \in \mathcal{X}}$ .

Then, after learning the outcome  $x_1$ , Bob chooses a new set of generalized measurement operators  $\{M_x^{(2|x_1)}\}_{x \in \mathcal{X}}$ . Note that this set of measurement operators can depend on  $x_1$ . Without loss of generality, we can keep the same outcome set  $\mathcal{X}$ , since Bob could set, e.g.  $M_{x_1}^{(2|x_1)} = 0$  to avoid guessing the same number twice. Bob measures his state and gets the outcome  $x_2$  with probability

$$p_{G_2|G_1X}(x_2|x_1, x) = \frac{1}{p_{G_1|X}(x_1|x)} \text{tr}[M_{x_2}^{(2|x_1)} M_{x_1}^{(1)} \rho_B^x M_{x_1}^{(1)\dagger} M_{x_2}^{(2|x_1)\dagger}]. \quad (6.24)$$

Multiplying by  $p_{G_1|X}(x_1|x)$  we see the joint distribution is given by

$$p_{G_1G_2|X}(x_1, x_2|x) = \text{tr}[M_{x_2}^{(2|x_1)} M_{x_1}^{(1)} \rho_B^x M_{x_1}^{(1)\dagger} M_{x_2}^{(2|x_1)\dagger}]. \quad (6.25)$$

To make his  $j$ th guess, we allow Bob to choose a new set of generalized measurement operators  $\{M_x^{(j|x_1, \dots, x_{j-1})}\}_{x \in \mathcal{X}}$  which may depend on the previous  $j - 1$  outcomes. Repeating the previous logic, in the end we find that

$$p_{G_1G_2 \dots G_K|X}(x_1, x_2, \dots, x_K|x) = \text{tr}[M_{x_K}^{(K|x_1, x_2, \dots, x_{K-1})} \dots M_{x_2}^{(2|x_1)} M_{x_1}^{(1)} \rho_B^x M_{x_1}^{(1)\dagger} M_{x_2}^{(2|x_1)\dagger} \dots M_{x_K}^{(K|x_1, x_2, \dots, x_{K-1})\dagger}]. \quad (6.26)$$

Under such a strategy, the possible random variables giving the number of guesses is given by



$$\mathcal{S}^{\text{Sequential}}(\rho_{XB}) := \left\{ N(\vec{G}, X) : (\vec{G}, X) \text{ satisfy (6.26) for some collections of measurement operators } \{M_{x_j}^{(j|x_1, x_2, \dots, x_{j-1})}\}_{x_j \in \mathcal{X}, j=1, \dots, K}, x_1, x_2, \dots, x_K \in \mathcal{X} \right\}. \quad (6.27)$$

**Theorem 6.3.1.** *Let  $\rho_{XB}$  be a c-q state as defined in (6.6) and  $K$  a natural number with  $K \leq |\mathcal{X}|$ . Then*

$$\mathcal{S}_K^{\text{Sequential}}(\rho_{XB}) = \mathcal{S}_K^{\text{Ordered}}(\rho_{XB}) = \mathcal{S}_K^{\text{Measured}}(\rho_{XB}). \quad (6.28)$$

Hence, all three sets of random variables of the number of guesses obtained from various classes of strategies all coincide. Hence, we call the single class that of *quantum strategies*, denoted  $\mathcal{S}_K^{\text{Quantum}}(\rho_{XB})$ .

*Proof.* The second equality was already stated in (6.21) and proven before that, and so it remains to prove the first equality. Consider a sequential strategy, with the notation of point 3 above. Define

$$E_{x_1, \dots, x_K} := M_{x_1}^{(1)\dagger} M_{x_2}^{(2|x_1)\dagger} \dots M_{x_K}^{(K|x_1, x_2, \dots, x_{K-1})\dagger} M_{x_K}^{(K|x_1, x_2, \dots, x_{K-1})} \dots M_{x_2}^{(2|x_1)} M_{x_1}^{(1)}. \quad (6.29)$$

We see that  $E_{x_1, \dots, x_K} = A^\dagger A$  for  $A = M_{x_K}^{(K|x_1, x_2, \dots, x_{K-1})} \dots M_{x_2}^{(2|x_1)} M_{x_1}^{(1)}$ , and hence is positive semidefinite. Moreover,

$$\sum_{x_1, \dots, x_K \in \mathcal{X}} E_{x_1, \dots, x_K} = I_B \quad (6.30)$$

as can be seen by first summing (6.29) over  $x_K$ , using

$$\sum_{x_K \in \mathcal{X}} M_{x_K}^{(K|x_1, x_2, \dots, x_{K-1})\dagger} M_{x_K}^{(K|x_1, x_2, \dots, x_{K-1})} = I_B \quad (6.31)$$

since  $\{M_x^{(K|x_1, x_2, \dots, x_{K-1})}\}_{x \in \mathcal{X}}$  is a POVM, and then similarly summing over  $x_{K-1}, x_{K-2}, \dots$ , and finally  $x_1$ . Let us write  $E_{\vec{x}}$  where  $\vec{x} = (x_1, \dots, x_K)$  for  $E_{x_1, \dots, x_K}$ . We have shown that  $\{E_{\vec{x}}\}_{\vec{x} \in \mathcal{X}^K}$  is a POVM. Moreover,

$$p_{G_1 G_2 \dots G_K | X}(x_1, x_2, \dots, x_K | x) = \text{tr}[E_{x_1, \dots, x_K} \rho_B^x]. \quad (6.32)$$

Hence, Bob's strategy is equivalent to simply performing the single POVM  $\{E_{\vec{x}}\}_{\vec{x} \in \mathcal{X}^K}$  once, obtaining an outcome  $\vec{x} = (x_1, \dots, x_K)$ , and then making  $x_1$  his first guess,  $x_2$

his second guess, and so forth. That is, any such strategy can be recast as an ordered strategy.

On the other hand, any such ordered strategy can be reformulated as an adaptive strategy, by the following recursive approach. Suppose that we are given  $\{E_{\vec{y}}\}_{\vec{y} \in \mathcal{X}^K}$ . For each  $x_1 \in \mathcal{X}$ , define

$$M_{x_1}^{(1)} = \sqrt{\sum_{x_2, \dots, x_K \in \mathcal{X}} E_{x_1, \dots, x_K}} \quad (6.33)$$

where we have chosen the positive semidefinite square root. We have that

$$\sum_{x_1 \in \mathcal{X}} M_{x_1}^{(1)\dagger} M_{x_1}^{(1)} = \sum_{x_1 \in \mathcal{X}} (M_{x_1}^{(1)})^2 = \sum_{x_1 \in \mathcal{X}} \sum_{x_2, \dots, x_K \in \mathcal{X}} E_{x_1, \dots, x_K} = I_B, \quad (6.34)$$

so  $\{M_{x_1}^{(1)}\}_{x_1 \in \mathcal{X}}$  is indeed a POVM with outcomes in  $\mathcal{X}$ . Next, for each  $x_1 \in \mathcal{X}$ , corresponding to obtaining outcome  $x_1$  on the first measurement, we define a POVM  $\{M_{x_2}^{(2|x_1)}\}_{x_2 \in \mathcal{X}}$  by

$$M_{x_2}^{(2|x_1)} = \sqrt{(M_{x_1}^{(1)})^{-1} \sum_{x_3, \dots, x_K \in \mathcal{X}} E_{x_1, \dots, x_K} (M_{x_1}^{(1)})^{-1}}. \quad (6.35)$$

Then

$$\sum_{x_2 \in \mathcal{X}} (M_{x_2}^{(2|x_1)})^2 = (M_{x_1}^{(1)})^{-1} \sum_{x_2 \in \mathcal{X}} \sum_{x_3, \dots, x_K \in \mathcal{X}} E_{x_1, \dots, x_K} (M_{x_1}^{(1)})^{-1} \quad (6.36)$$

$$= (M_{x_1}^{(1)})^{-1} (M_{x_1}^{(1)})^2 (M_{x_1}^{(1)})^{-1} = I_B. \quad (6.37)$$

Likewise, we define

$$M_{x_3}^{(3|x_1, x_2)} = \sqrt{(M_{x_2}^{(2|x_1)})^{-1} (M_{x_1}^{(1)})^{-1} \sum_{x_4, \dots, x_K \in \mathcal{X}} E_{x_1, \dots, x_K} (M_{x_1}^{(1)})^{-1} (M_{x_2}^{(2|x_1)})^{-1}}. \quad (6.38)$$

Then

$$\sum_{x_3 \in \mathcal{X}} (M_{x_3}^{(3|x_1, x_2)})^2 = (M_{x_2}^{(2|x_1)})^{-1} (M_{x_1}^{(1)})^{-1} \sum_{x_3 \in \mathcal{X}} \sum_{x_4, \dots, x_K \in \mathcal{X}} E_{x_1, \dots, x_K} (M_{x_1}^{(1)})^{-1} (M_{x_2}^{(2|x_1)})^{-1} \quad (6.39)$$

$$= (M_{x_2}^{(2|x_1)})^{-1} (M_{x_2}^{(2|x_1)})^2 (M_{x_2}^{(2|x_1)})^{-1} = I_B. \quad (6.40)$$

Repeating this process, we define

$$M_{x_\ell}^{(\ell|x_1, x_2, \dots, x_{\ell-1})} = \sqrt{(M_{x_{\ell-1}}^{(\ell-1|x_1, x_2, \dots, x_{\ell-2})})^{-1} \dots (M_{x_1}^{(1)})^{-1} \sum_{x_{\ell+1}, \dots, x_K \in \mathcal{X}} E_{x_1, \dots, x_K} (M_{x_1}^{(1)})^{-1} \dots M_{x_{\ell-1}}^{(\ell-1|x_1, x_2, \dots, x_{\ell-2})})^{-1}} \quad (6.41)$$

to obtain a POVM for step  $\ell$  (to use when having obtained outcomes  $x_1, \dots, x_{\ell-1}$  during the previous steps). At the last step,  $\ell = K$ , there is no sum, namely

$$M_{x_K}^{(K|x_1, x_2, \dots, x_{K-1})} = \sqrt{(M_{x_{K-1}}^{(K-1|x_1, x_2, \dots, x_{K-2})})^{-1} \dots (M_{x_1}^{(1)})^{-1} E_{x_1, \dots, x_K} (M_{x_1}^{(1)})^{-1} \dots M_{x_{K-1}}^{(K-1|x_1, x_2, \dots, x_{K-2})})^{-1}}. \quad (6.42)$$

Lastly, we check that by design, (6.29) holds. Thus, we can work backwards from that equation and see that our newly created adaptive strategy yields the same outcomes with the same probabilities as the initial ordered strategy.  $\square$

### Success metrics

Given a random variable  $X$  and a maximal number  $K$  of allowed guesses, how do we measure the success of a guessing strategy  $\vec{G}$ ? We will focus on expectations of  $N(\vec{G}, X)$ . In particular, we consider the expected number of guesses required to guess correctly:

$$\mathbb{E}[N(\vec{G}, X)] = \begin{cases} \sum_{k=1}^K k \cdot p_{N(\vec{G}, X)}(k) & \text{if } p_{N(\vec{G}, X)}(\infty) = 0 \\ \infty & \text{if } p_{N(\vec{G}, X)}(\infty) > 0. \end{cases} \quad (6.43)$$

As with the generalized guesswork without side information defined in (6.4), we consider a cost vector  $\vec{c} \in (\mathbb{R} \cup \{\infty\})^{|\mathcal{X}|}$  with

$$0 \leq c_1 \leq c_2 \leq \dots \leq c_{|\mathcal{X}|}. \quad (6.44)$$

Then we define the modified expectation

$$E_{\vec{c}}(N(\vec{G}, X)) := \sum_{k=1}^{|\mathcal{X}|} c_k \cdot p_{N(\vec{G}, X)}(k).$$

Imposing a maximal number  $K < |\mathcal{X}|$  of allowed guesses is equivalent to choosing  $c_{K+1} = \dots = c_{|\mathcal{X}|} = \infty$ , using the convention  $\infty \cdot 0 = 0$ . Accordingly, we implicitly associate  $K$  with  $\vec{c}$  in all the following via the rule that  $K = |\mathcal{X}|$  if and only if  $c_{|\mathcal{X}|} < \infty$ , and otherwise  $K = \min\{i : c_i = \infty\}$ . The case  $K = |\mathcal{X}|$  therefore corresponds to  $|\mathcal{X}|$  guesses being allowed, each with finite cost, and the case  $K < |\mathcal{X}|$  corresponds to a limited number of allowed guesses, with a corresponding infinite cost if the correct answer is not obtained in  $K$  guesses.

Given a c-q state  $\rho_{XB}$  and a cost vector  $\vec{c}$  as in (6.44), we define the generalized guesswork with quantum side information as

$$G_{\vec{c}}(X|B)_{\rho} := \inf_{N \in \mathcal{S}_K^{\text{Quantum}}(\rho_{XB})} E_{\vec{c}}(N). \quad (6.45)$$

Likewise, given a joint distribution  $p_{XY}$ , let

$$G_{\vec{c}}(X|Y)_p := \inf_{N \in \mathcal{S}_K^{\text{Classical}}(p_{XY})} E_{\vec{c}}(N). \quad (6.46)$$

From the equality

$$\mathcal{S}_K^{\text{Quantum}}(\rho_{XB}, K) = \mathcal{S}_K^{\text{Measured}}(\rho_{XB}) \quad (6.47)$$

of Theorem 6.3.1 it follows that

$$G_{\vec{c}}(X|B)_{\rho} = \inf_{\{E_y\}_{y \in \mathcal{Y}}} G_{\vec{c}}(X|Y)_p \quad (6.48)$$

where the infimum is over all finite alphabets  $\mathcal{Y}$  and POVMs  $\{E_y\}_{y \in \mathcal{Y}}$  and  $p_{XY}(x, y) = p_X(x) \text{tr}[E_y \rho_B^x]$ .

In the standard case in which  $\vec{c} = (1, 2, \dots, |\mathcal{X}|)$ , we define the *guesswork with quantum side information* as

$$G(X|B) \equiv G(X|B)_{\rho} := G_{\vec{c}}(X|B)_{\rho} \quad (6.49)$$

and likewise define  $G(X|Y)_p = G_{\vec{c}}(X|Y)_p$  in the case of classical side information  $Y$ .

*Remark.* In the work [CCWF15], guesswork with quantum side information was defined by the right-hand side of (6.48) (with  $\vec{c} = (1, 2, \dots, |\mathcal{X}|)$ ). Moreover, Proposition 1 of that work shows that the infimum in (6.48) in that case may be restricted to POVMs whose elements are all rank one.

### 6.3.3 A semidefinite optimization representation

The task of calculating  $G_{\vec{c}}(X|B)_\rho$  as defined in (6.45) can be written as a semidefinite optimization problem, as was found in [CCWF15]. In this section, we present a different derivation of that fact yielding in (6.58) a representation dual to the one found in [CCWF15]. This representation is used in Section 6.3.4 to prove that the guesswork  $G_{\vec{c}}(X|B)_\rho$  is a concave and Lipschitz-continuous function of the c-q state  $\rho_{XB}$  (if  $K = |\mathcal{X}|$ ). Section 6.3.5 presents other optimization formulations of the guesswork with the aim of computing the quantity.

Consider an ordered strategy  $\vec{G}$  with a set of POVMs  $\{E_{\vec{g}}\}_{\vec{g} \in \mathcal{X}^K}$ . Then since  $p_{\vec{G},X}(\vec{g}, x) = p_X(x) \text{tr}[E_{\vec{g}}\rho_B^x]$ , we have

$$c_k p_{N(\vec{G}, X)}(k) = c_k \sum_{x \in \mathcal{X}} p_X(x) \sum_{\substack{\vec{g} \in \mathcal{X}^K \\ N(\vec{g}, x) = k}} \text{tr}[E_{\vec{g}}\rho_B^x] \quad (6.50)$$

and hence if

$$\sum_{x \in \mathcal{X}} p_X(x) \sum_{\substack{\vec{g} \in \mathcal{X}^K \\ N(\vec{g}, x) = \infty}} \text{tr}[E_{\vec{g}}\rho_B^x] > 0 \quad (6.51)$$

then  $E_{\vec{c}}(N(\vec{G}, X)) = \infty$ , and otherwise

$$E_{\vec{c}}(N(\vec{G}, X)) = \sum_{k=1}^K c_k \sum_{x \in \mathcal{X}} p_X(x) \sum_{\substack{\vec{g} \in \mathcal{X}^K \\ N(\vec{g}, x) = k}} \text{tr}[E_{\vec{g}}\rho_B^x] \quad (6.52)$$

$$= \sum_{\vec{g} \in \mathcal{X}^K} \sum_{x \in \mathcal{X}} c_{N(\vec{g}, x)} p_X(x) \text{tr}[E_{\vec{g}}\rho_B^x] \quad (6.53)$$

$$= \sum_{\vec{g} \in \mathcal{X}^K} \text{tr}[R_{\vec{g}} E_{\vec{g}}] \quad (6.54)$$

where we define  $R_{\vec{g}} := \sum_{x \in \mathcal{X}} p_X(x) c_{N(\vec{g}, x)} \rho_B^x$  for  $\vec{g} \in \mathcal{X}^K$ . Note that if  $K = |\mathcal{X}|$  then (6.51) does not hold.

Thus,

$$\begin{aligned}
G_{\vec{c}}(X|B)_\rho = \quad & \text{minimize} \quad \sum_{\vec{g} \in \mathcal{X}^K} \text{tr}[R_{\vec{g}} E_{\vec{g}}] \\
\text{subject to} \quad & E_{\vec{g}} \geq 0 \quad \forall \vec{g} \in \mathcal{X}^K \\
& \sum_{\vec{g} \in \mathcal{X}^K} E_{\vec{g}} = \mathbf{1}_B \\
& \sum_{x \in \mathcal{X}} p_X(x) \sum_{\substack{\vec{g} \in \mathcal{X}^K \\ N(\vec{g}, x) = \infty}} \text{tr}[E_{\vec{g}} \rho_B^x] = 0
\end{aligned} \tag{6.55}$$

and where an infeasible problem is associated to the value  $+\infty$ . Note that in the case  $K = |\mathcal{X}|$  the last constraint is trivially satisfied; in this case, the problem is always feasible. The expression in (6.55) clarifies that  $R_{\vec{g}}$  has an interpretation as a cost operator corresponding to the guessing outcome  $\vec{g}$ . Since  $\sum_{\vec{g} \in \mathcal{X}^K} \text{tr}[R_{\vec{g}} E_{\vec{g}}]$  is linear in each positive semidefinite (matrix) variable  $E_{\vec{g}}$ , (6.55) gives an SDP representation of  $G_{\vec{c}}(X|B)_\rho$ . This program has  $|\mathcal{X}|^K$  variables (each  $d_B \times d_B$  complex positive semidefinite matrices), subject to two constraints. Note, however, since the cost vector  $\vec{c}$  is increasing, any guess  $\vec{h} \in \mathcal{X}^K$  with repeated elements is a suboptimal guessing order, in the sense that if  $\{E_{\vec{g}}\}_{\vec{g} \in \mathcal{X}^K}$  is a POVM with  $E_{\vec{h}} \neq 0$ , and  $\vec{h}' \in \mathcal{X}^K$  only differs from  $\vec{h}$  by replacing repeated elements such that  $\vec{h}'$  has no repeated elements, then the POVM defined by

$$\tilde{E}_{\vec{g}} := \begin{cases} E_{\vec{g}} & \vec{g} \neq \vec{h} \text{ and } \vec{g} \neq \vec{h}' \\ 0 & \vec{g} = \vec{h} \\ E_{\vec{h}} + E_{\vec{h}'} & \vec{g} = \vec{h}' \end{cases} \tag{6.56}$$

has  $\sum_{\vec{g} \in \mathcal{X}^K} \text{tr}[R_{\vec{g}} \tilde{E}_{\vec{g}}] \leq \sum_{\vec{g} \in \mathcal{X}^K} \text{tr}[R_{\vec{g}} E_{\vec{g}}]$ . Hence, we may restrict to the outcome space

$$\mathcal{X}_{\neq}^K := \{\vec{g} \in \mathcal{X}^K : g_i \neq g_j, \forall i \neq j\} \subseteq \mathcal{X}^K. \tag{6.57}$$

Note  $|\mathcal{X}_{\neq}^K| = \frac{|\mathcal{X}|!}{(|\mathcal{X}|-K)!}$ , and in the case in which  $K = |\mathcal{X}|$ , the set  $\mathcal{X}_K$  is just the set of permutations of  $\mathcal{X}$ . Hence, (6.55) can be re-written as the following smaller problem:

$$\begin{aligned}
G_{\bar{c}}(X|B)_{\rho} = & \text{minimize} && \sum_{\vec{g} \in \mathcal{X}_{\neq}^K} \text{tr}[R_{\vec{g}} E_{\vec{g}}] \\
& \text{subject to} && E_{\vec{g}} \geq 0 \quad \forall \vec{g} \in \mathcal{X}_{\neq}^K \\
& && \sum_{\vec{g} \in \mathcal{X}_{\neq}^K} E_{\vec{g}} = \mathbf{1}_B \\
& && \sum_{x \in \mathcal{X}} p_X(x) \sum_{\substack{\vec{g} \in \mathcal{X}_{\neq}^K \\ N(\vec{g}, x) = \infty}} \text{tr}[E_{\vec{g}} \rho_B^x] = 0
\end{aligned} \tag{6.58}$$

Note that in the case  $K < |\mathcal{X}|$ , there exists a feasible point, and hence a solution, if and only if there exists a POVM  $\{E_{\vec{g}}\}_{\vec{g} \in \mathcal{X}_{\neq}^K}$  such that for all  $x \in \mathcal{X}$  and  $\vec{g} \in \mathcal{X}_{\neq}^K$  with  $x \notin \vec{g}$ , we have  $\text{tr}[E_{\vec{g}} \rho_B^x] = 0$ . Whether or not this holds a priori depends on the particular state  $\rho_{XB}$ . However, when  $K = |\mathcal{X}|$ , the final constraint can be removed and for any state  $\rho_{XB}$ , the problem (6.58) trivially has a feasible point (e.g.  $E_{\vec{g}} = \frac{1}{|\mathcal{X}_{\neq}^K|} \mathbf{1}_X$ ). In the following, we restrict to the case  $K = |\mathcal{X}|$  for simplicity.

*Remark.* This optimization problem has the same form as that of discriminating quantum states in an ensemble, as described in, e.g., [Wat18, Section 3.2.1]. Note, however, that (1) the  $R_{\vec{g}}$  are positive semidefinite but not normalized, and (2) the case of having two copies of the unknown state, in the guessing framework, does not correspond to  $R_{\vec{g}}^{\otimes 2}$ . Nevertheless, slight modifications to [Wat18, Theorem 3.9] show that a POVM  $\{E_{\vec{g}}\}_{\vec{g} \in \mathcal{X}_{\neq}^K}$  is optimal for (6.58) if and only if

$$Y = \sum_{\vec{g} \in \mathcal{X}_{\neq}^K} R_{\vec{g}} E_{\vec{g}} \tag{6.59}$$

satisfies  $Y \leq R_{\vec{g}}$  for all  $\vec{g} \in \mathcal{X}_{\neq}^K$ .

*Remark.* The set of POVMs is convex and since the objective function is linear, any minimizer for (6.58) may be decomposed into extremal POVMs which are also minimizers. By [Par99, Corollary 2.2], for any extremal POVM on a Hilbert space of size  $d_B$  has at most  $d_B^2$  non-zero elements. Hence, there exist minimizers of (6.58) with at most  $d_B^2$  non-zero elements (even though  $|\mathcal{X}_{\neq}^K|$  could be far larger than  $d_B^2$ ). Let  $S \subseteq \mathcal{X}_{\neq}^K$  be a set of  $d_B^2$  points such that there exists  $\{\tilde{E}_{\vec{g}}\}_{\vec{g} \in S}$  with  $\tilde{E}_{\vec{g}} \geq 0$ ,  $\sum_{\vec{g} \in S} \tilde{E}_{\vec{g}} = \mathbf{1}_B$ , and

$$G_{\bar{c}}(X|B)_{\rho} = \sum_{\vec{g} \in S} \text{tr}[\tilde{E}_{\vec{g}} R_{\vec{g}}]. \tag{6.60}$$

Then (6.58) holds with  $\mathcal{X}_{\neq}^K$  replaced by  $S$ , namely

$$\begin{aligned} G_{\vec{c}}(X|B)_{\rho} = & \text{minimize} \quad \sum_{\vec{g} \in S} \text{tr}[R_{\vec{g}} E_{\vec{g}}] \\ & \text{subject to} \quad E_{\vec{g}} \geq 0 \quad \forall \vec{g} \in S \\ & \quad \sum_{\vec{g} \in S} E_{\vec{g}} = \mathbf{1}_B. \end{aligned} \quad (6.61)$$

Note the “ $\leq$ ” direction of the equality (6.61) is trivial, since given a minimizer  $\{E_{\vec{g}}\}_{\vec{g} \in S}$  for (6.61), simply extending it by choosing  $E_{\vec{g}} = 0$  for  $\vec{g} \notin S$  gives a feasible point for the optimization problem on the right-hand side of (6.58). The “ $\geq$ ” direction follows from the existence of the  $\{\tilde{E}_{\vec{g}}\}_{\vec{g} \in S}$  described above. While this produces a much smaller problem, the set  $S$  depends in general on  $\rho$  and the  $\{R_{\vec{g}}\}_{\vec{g} \in \mathcal{X}_{\neq}^K}$ , and the task of finding  $S$  is as difficult as solving the original problem.

### 6.3.4 Concavity and continuity of the guesswork with quantum side information

**Proposition 6.3.2.** *For each finite cost vector  $\vec{c}$  satisfying (6.44), the function*

$$\rho_{XB} \mapsto G_{\vec{c}}(X|B)_{\rho} \quad (6.62)$$

*from the set of c-q states of the form (6.6) to  $\mathbb{R}_{\geq 0}$ , is concave.*

*Proof.* Recall  $c_{|\mathcal{X}|} < \infty$  means  $K = |\mathcal{X}|$ . For  $\vec{g} \in \mathcal{X}_{\neq}^K$ , and  $\rho_{XB}$  a c-q state, the quantity  $R_{\vec{g}}^{\rho} := \sum_{x \in \mathcal{X}} p_X(x) c_{N(\vec{g}, x)} \rho_B^x$  can be expressed as

$$R_{\vec{g}}^{\rho} = \text{tr}_X \left[ \left( \sum_{x \in \mathcal{X}} c_{N(\vec{g}, x)} |x\rangle \langle x|_X \otimes I_B \right) \rho_{XB} \right] \quad (6.63)$$

and hence is linear in  $\rho_{XB}$ . Then for each POVM  $(E_{\vec{g}})_{\vec{g} \in \mathcal{X}_{\neq}^K}$ ,

$$\rho_{XB} \mapsto \sum_{\vec{g} \in \mathcal{X}_{\neq}^K} \text{tr}[R_{\vec{g}}^{\rho} E_{\vec{g}}] \quad (6.64)$$



is linear in  $\rho_{XB}$ . The arbitrary infimum of concave functions, and in particular linear functions, is concave, and hence

$$G_{\vec{c}}(X|B)_\rho \equiv \min_{(E_{\vec{g}})_{\vec{g} \in \mathcal{X}_{\neq}^K}} \sum_{\vec{g} \in \mathcal{X}_{\neq}^K} \text{tr}[R_{\vec{g}}^\rho E_{\vec{g}}], \quad (6.65)$$

where the minimum is taken over all POVMs on system  $B$  with outcomes in  $\mathcal{X}_{\neq}^K$ , is concave.  $\square$

**Proposition 6.3.3.** *For each finite cost vector  $\vec{c}$  satisfying (6.44), the function*

$$\rho_{XB} \mapsto G_{\vec{c}}(X|B)_\rho \quad (6.66)$$

*from the set of  $c$ - $q$  states of the form (6.6) to  $\mathbb{R}_{\geq 0}$ , is Lipschitz continuous, satisfying the bound*

$$|G_{\vec{c}}(X|B)_\rho - G_{\vec{c}}(X|Y)_\sigma| \leq 2c_{|\mathcal{X}|} T(\rho_{XB}, \sigma_{XB}). \quad (6.67)$$

*Moreover, the optimal Lipschitz constant  $k$  satisfies*

$$c_{|\mathcal{X}|} - c_1 \leq k \leq 2c_{|\mathcal{X}|}, \quad (6.68)$$

*and the lower bound is tight: there exist  $\rho_{XB}$  and  $\sigma_{XB}$  such that  $|G_{\vec{c}}(X|B)_\rho - G_{\vec{c}}(X|Y)_\sigma| = c_{|\mathcal{X}|} - c_1$ .*

*Remark.* In the standard case case  $\vec{c} = (1, \dots, |\mathcal{X}|)$ , the optimal Lipschitz constant for  $G_{\vec{c}}(X|B)$  thus satisfies

$$|\mathcal{X}| - 1 \leq k \leq 2|\mathcal{X}|,$$

and hence the bound of Proposition 6.3.3 is tight up to constant factors in this case. An interesting open question is whether or not equality is achieved in the lower bound of (6.68).

*Proof.* To establish the lower bound of (6.68), note that product states  $\rho_{XB} = \omega_X \otimes \sigma_B$  satisfy  $G_{\vec{c}}(X|B)_\rho = G_{\vec{c}}(X)_\omega$ . Hence, Proposition 6.2.1 shows that  $k \geq c_{|\mathcal{X}|} - c_1$ , and the tightness of the bound follows from the tightness of Proposition 6.2.1.

Next, let us establish the upper bound. Define

$$f(\rho_{XB}, \{E_{\vec{g}}\}_{\vec{g} \in \mathcal{X}_{\neq}^K}) := \sum_{\vec{g} \in \mathcal{X}_{\neq}^K} \text{tr}[R_{\vec{g}}^\rho E_{\vec{g}}]. \quad (6.69)$$

Then, by linearity (as discussed in the proof of Proposition 6.3.2),

$$f(\rho_{XB}, \{E_{\vec{g}}\}_{\vec{g} \in \mathcal{X}_{\neq}^K}) - f(\sigma_{XB}, \{E_{\vec{g}}\}_{\vec{g} \in \mathcal{X}_{\neq}^K}) = \sum_{\vec{g} \in \mathcal{X}_{\neq}^K} \text{tr}[R_{\vec{g}}^{\rho-\sigma} E_{\vec{g}}] \quad (6.70)$$

$$= \sum_{\vec{g} \in \mathcal{X}_{\neq}^K} \text{tr}[\text{tr}_X[C_{XB}^{(\vec{g})} \Delta_{XB}] E_{\vec{g}}] \quad (6.71)$$

using (6.63), where  $C_{XB}^{(\vec{g})} := \sum_{x \in \mathcal{X}} c_{N(\vec{g}, x)} |x\rangle\langle x| \otimes I_B \geq 0$  and  $\Delta_{XB} := \rho_{XB} - \sigma_{XB}$ . Since  $C_{XB}^{(\vec{g})}$  and  $\Delta_{XB}$  commute, using the c-q structure of each, we have

$$f(\rho_{XB}, \{E_{\vec{g}}\}_{\vec{g} \in \mathcal{X}_{\neq}^K}) - f(\sigma_{XB}, \{E_{\vec{g}}\}_{\vec{g} \in \mathcal{X}_{\neq}^K}) = \sum_{\vec{g} \in \mathcal{X}_{\neq}^K} \text{tr}[C_{XB}^{(\vec{g})} \Delta_{XB} (I_X \otimes E_{\vec{g}})] \quad (6.72)$$

$$= \text{tr} \left[ \Delta_{XB} \sum_{\vec{g} \in \mathcal{X}_{\neq}^K} C_{XB}^{(\vec{g})} (I_X \otimes E_{\vec{g}}) \right]. \quad (6.73)$$

Set

$$F_{XB} := \sum_{\vec{g} \in \mathcal{X}_{\neq}^K} C_{XB}^{(\vec{g})} (I_X \otimes E_{\vec{g}}) = \sum_{x \in \mathcal{X}} \sum_{\vec{g} \in \mathcal{X}_{\neq}^K} c_{N(\vec{g}, x)} |x\rangle\langle x| \otimes E_{\vec{g}}. \quad (6.74)$$

Since  $c_{N(\vec{g}, x)} \leq c_{|\mathcal{X}|}$  for each  $x \in \mathcal{X}$  and  $\vec{g} \in \mathcal{X}_{\neq}^K$ , we have that  $F_{XB} \leq c_{|\mathcal{X}|} \sum_{x \in \mathcal{X}} \sum_{\vec{g} \in \mathcal{X}_{\neq}^K} |x\rangle\langle x| \otimes E_{\vec{g}}$  in semidefinite order. Performing the sums, we have  $F_{XB} \leq c_{|\mathcal{X}|} I_X \otimes I_B$  and hence  $\|F_{XB}\|_{\infty} \leq c_{|\mathcal{X}|}$ . Thus,

$$f(\rho_{XB}, \{E_{\vec{g}}\}_{\vec{g} \in \mathcal{X}_{\neq}^K}) - f(\sigma_{XB}, \{E_{\vec{g}}\}_{\vec{g} \in \mathcal{X}_{\neq}^K}) = \text{tr}[\Delta_{XB} F_{XB}] \quad (6.75)$$

$$\leq \|\Delta_{XB} F_{XB}\|_1 \quad (6.76)$$

$$\leq \|\Delta_{XB}\|_1 \|F_{XB}\|_{\infty} \quad (6.77)$$

$$\leq c_{|\mathcal{X}|} \|\rho_{XB} - \sigma_{XB}\|_1 \quad (6.78)$$

$$= 2c_{|\mathcal{X}|} T(\rho_{XB}, \sigma_{XB}). \quad (6.79)$$

using Hölder's inequality in the third to last inequality. Swapping the role of  $\rho_{XB}$  and  $\sigma_{XB}$  completes the proof.  $\square$

This completes the main goal of this chapter: establishing the Lipschitz continuity of the guesswork with quantum side information. In the remainder of the chapter, we will look at some further aspects of the guesswork with quantum side information in order to establish methods for computing it (or computing bounds of it), and to obtain entropic bounds on the quantity.

### 6.3.5 Computing the guesswork

The SDP formulation of the guesswork described in (6.58) provides the first method of computing the quantity: by solving the semidefinite optimization problem. The difficulty in this approach is that the number of variables in the problem scales as  $|\mathcal{X}_{\neq}^K| = \frac{|\mathcal{X}|!}{(|\mathcal{X}|-K)!}$ . For  $K$  scaling with  $|\mathcal{X}|$  (e.g.,  $K = |\mathcal{X}|$ ), the number of variables grows extremely quickly. In this section, we present two alternatives in the case  $K = |\mathcal{X}|$ : an algorithm using an active-set method to establish an upper bound, and a reformulation of the problem as a *mixed-integer SDP* with polynomial dependence on  $|\mathcal{X}|$ .

#### The dual problem

First, we compute the dual problem to (6.58), in the case  $K = |\mathcal{X}|$ . Consider the Lagrangian

$$\mathcal{L}((E_{\vec{g}})_{\vec{g} \in \mathcal{X}_{\neq}^K}, (\lambda_{\vec{g}})_{\vec{g} \in \mathcal{X}_{\neq}^K}, \nu) = \sum_{\vec{g} \in \mathcal{X}_{\neq}^K} \langle R_{\vec{g}}, E_{\vec{g}} \rangle - \sum_{\vec{g} \in \mathcal{X}_{\neq}^K} \langle \lambda_{\vec{g}}, E_{\vec{g}} \rangle + \left\langle \nu, \sum_{\vec{g} \in \mathcal{X}_{\neq}^K} E_{\vec{g}} - \mathbf{1}_B \right\rangle \quad (6.80)$$

$$= \sum_{\vec{g} \in \mathcal{X}_{\neq}^K} \langle R_{\vec{g}} - \lambda_{\vec{g}} + \nu, E_{\vec{g}} \rangle - \text{tr}[\nu] \quad (6.81)$$

where we have introduced the Hilbert–Schmidt inner product  $\langle A, B \rangle = \text{tr}[A^\dagger B]$ , and where  $\lambda_{\vec{g}} \geq 0$  is the dual variable to the inequality constraint  $E_{\vec{g}} \geq 0$ , and  $\nu = \nu^\dagger$  is the dual variable to the equality constraint  $\sum_{\vec{g} \in \mathcal{X}_{\neq}^K} E_{\vec{g}} = \mathbf{1}_B$ . As shown in, e.g., [BV04], the primal problem (6.58) can be expressed as

$$\min_{(E_{\vec{g}})_{\vec{g} \in \mathcal{X}_{\neq}^K}} \max_{\lambda_{\vec{g}} \geq 0, \nu} \mathcal{L}((E_{\vec{g}})_{\vec{g} \in \mathcal{X}_{\neq}^K}, (\lambda_{\vec{g}})_{\vec{g} \in \mathcal{X}_{\neq}^K}, \nu) \quad (6.82)$$

while the dual problem is given by

$$\max_{\lambda_{\vec{g}} \geq 0, \nu} \min_{(E_{\vec{g}})_{\vec{g} \in \mathcal{X}_{\neq}^K}} \mathcal{L}((E_{\vec{g}})_{\vec{g} \in \mathcal{X}_{\neq}^K}, (\lambda_{\vec{g}})_{\vec{g} \in \mathcal{X}_{\neq}^K}, \nu). \quad (6.83)$$

If  $R_{\vec{g}} - \lambda_{\vec{g}} + \nu \neq 0$  for any  $\vec{g} \in \mathcal{X}_{\neq}^K$ , then the inner minimization in (6.83) yields  $-\infty$ . Hence,

$$\min_{(E_{\vec{g}})_{\vec{g} \in \mathcal{X}_{\neq}^K}} \mathcal{L}((E_{\vec{g}})_{\vec{g} \in \mathcal{X}_{\neq}^K}, (\lambda_{\vec{g}})_{\vec{g} \in \mathcal{X}_{\neq}^K}, \nu) = \begin{cases} -\infty & R_{\vec{g}} - \lambda_{\vec{g}} + \nu \neq 0 \quad \exists \vec{g} \in \mathcal{X}_{\neq}^K \\ -\text{tr}[\nu] & \text{else.} \end{cases} \quad (6.84)$$

The constraint  $\lambda_{\vec{g}} \geq 0$  and  $R_{\vec{g}} - \lambda_{\vec{g}} + \nu = 0$  imply the semidefinite inequality  $-\nu \leq R_{\vec{g}}$ . Writing  $Y = -\nu$  and maximizing over  $\lambda_{\vec{g}} \geq 0$ , (6.83) becomes

$$\begin{aligned} & \text{maximize} && \text{tr}[Y] \\ & \text{subject to} && Y = Y^\dagger \\ & && Y \leq R_{\vec{g}} \quad \forall \vec{g} \in \mathcal{X}_{\neq}^K \end{aligned} \tag{6.85}$$

Since (6.58) is strictly feasible (e.g.,  $E_{\vec{g}} = \mathbf{1}_B \frac{1}{|\mathcal{X}_{\neq}^K|}$  is a strictly feasible point), Slater's condition proves that strong duality holds. Hence, (6.85) obtains the same optimal value as (6.58). The formulation of the problem as given in (6.85) was previously found in the work [CCWF15, Proposition 3].

### A simple algorithm to compute upper bounds

The dual form of the SDP can be used to generate upper bounds on  $G_{\vec{c}}(X|B)_\rho$  simply by removing constraints. This provides an algorithm to find an upper bound on the objective function: Decide on some number of constraints  $\kappa$  to impose in total. Then,

1. Initialize an empty list  $L = \{\}$  corresponding to constraints to impose.
2. Set  $Y$  to be the identity matrix, as a first guess at the optimal dual variable.
3. If  $Y$  satisfies  $Y \leq R_{\vec{g}}$  for all  $\vec{g} \in \mathcal{X}_{\neq}^K$ , then  $Y$  is the maximizer of the dual problem (6.85), and the optimization is solved. Otherwise, find  $\vec{g} \in \mathcal{X}_{\neq}^K$  such that  $Y \not\leq R_{\vec{g}}$ , and add  $\vec{g}$  to the list  $L$ .
4. Solve the problem

$$\begin{aligned} & \text{maximize} && \text{tr}[Y] \\ & \text{subject to} && Y = Y^\dagger \\ & && Y \leq R_{\vec{g}} \quad \forall \vec{g} \in L \end{aligned} \tag{6.86}$$

and set  $Y$  to be its maximizer.

5. Repeat steps 2 and 3 until the list  $L$  has length  $\kappa$ .
6. Solve the problem one last time, and return the output.

In order to find a constraint that  $Y$  violates, a heuristic technique such as simulated annealing can be used. Moreover, in the case that there are too many constraints to fit into memory or check exhaustively, using an iterative technique (such as simulated annealing) is essential. If this algorithm was continued without imposing a limit on

the total number of constraints  $\kappa$  to impose, it would eventually yield the true value  $G_{\mathcal{C}}(X|B)_\rho$ . When a total number of constraints is limited, it provides an upper bound (since it is a relaxation of (6.85)). Note that since it takes exponential time to check if  $Y \leq R_{\vec{g}}$  for every  $\vec{g} \in \mathcal{X}_{\neq}^K$ , in practice  $\kappa$  must be chosen low enough so that violated constraints can be found quickly enough (or  $\kappa$  can be chosen adaptively by adding constraints until finding a new violated constraint violates some predetermined time limit).

However, even with a limit  $\kappa$  on the total number of constraints, this algorithm can in theory yield the true value  $G_{\mathcal{C}}(X|B)_\rho$ . Note that the dual problem to (6.61) is

$$\begin{aligned} & \text{maximize} && \text{tr}[Y] \\ & \text{subject to} && Y = Y^\dagger \\ & && Y \leq R_{\vec{g}} \quad \forall \vec{g} \in S \end{aligned} \tag{6.87}$$

where  $S \subseteq \mathcal{X}_{\neq}^K$  has  $|S| = d_B^2$  and is described in the remark above. Hence, if  $L$  in (6.86) equals  $S$ , then the algorithm finds the true value  $G_{\mathcal{C}}(X|B)_\rho$ , not just an upper bound. Thus,  $\kappa = d_B^2$  suffices if the constraints  $\vec{g}$  can be chosen precisely to obtain  $L = S$ . In general, finding  $S$  is as difficult as solving the original problem. Nonetheless, this motivates why choosing a relatively small value of  $\kappa$  (such as  $d_B^2$ ) can still yield good upper bounds.

### A mixed-integer reformulation

The problem can be formulated another way as a mixed-integer SDP, i.e. an SDP that has additional integer or binary constraints. Consider a POVM  $\{F_j\}_{j=1}^M$  with  $M$  outcomes. When outcome  $j$  is obtained, Bob guesses in some order  $\vec{g}^{(j)} \in \mathcal{X}_{\neq}^K$ . Then consider the problem

$$\begin{aligned} & \text{minimize} && \sum_{x \in \mathcal{X}, j=1, \dots, M} p_X(x) c_{N(\vec{g}^{(j)}, x)} \text{tr}[F_j \rho_B^x] \\ & \text{subject to} && F_j \geq 0 \quad j = 1, \dots, M, \\ & && \vec{g}^{(j)} \in \mathcal{X}_{\neq}^K, j = 1, \dots, M, \\ & && \sum_{j=1}^M F_j = \mathbf{1}_B. \end{aligned} \tag{6.88}$$

This optimization is not an SDP, since the dependence on the optimization variables  $\{\vec{g}^{(j)}\}_{j=1}^M$  and  $\{F_j\}$  is not linear, and  $\vec{g}^{(j)} \in \mathcal{X}_{\neq}^K$  is a discrete constraint. In the case that  $K = |\mathcal{X}|$ , we will be able to remove the nonlinearity, although not the discrete

variables. This yields a *mixed-integer* SDP: an optimization problem such that if all integral constraints were removed, the result would be an SDP. We proceed as follows.

Under the condition  $K = |\mathcal{X}|$ , we may restrict to considering guessing orders that are permutations without loss of generality; other guessing orders have repeated guesses, which can only increase the objective function. In this case, the outcome  $\infty$  never occurs, and for each  $\vec{g} \in S_{|\mathcal{X}|}$ , the quantity  $(c_{N(\vec{g},x)})_{x \in \mathcal{X}}$  satisfies

$$(c_{N(\vec{g},x)})_{x \in \mathcal{X}} = \vec{g}^{-1}(c), \quad (6.89)$$

where  $\vec{g}^{-1}$  is the inverse permutation to  $\vec{g}$ , and  $c$  is the cost vector. Here,  $S_n$  is the set of permutations on  $\{1, \dots, n\}$ . Let  $P^{(j)}$  be an  $|\mathcal{X}| \times |\mathcal{X}|$  matrix representation of the permutation  $\vec{g}^{(j)-1}$ . Then  $(P^{(j)}c)_x = \sum_{y \in \mathcal{X}} P_{xy}^{(j)} c_y = c_{N(\vec{g},x)}$ . Hence, the optimization (6.88) can be reformulated as

$$\begin{aligned} & \text{minimize} && \sum_{x,y \in \mathcal{X}, j=1, \dots, M} p_X(x) P_{xy}^{(j)} c_y \operatorname{tr}[F_j \rho_B^x] \\ & \text{subject to} && F_j \in \mathcal{M}_{d_B} \quad \forall j \in [M], \\ & && P_{xy}^{(j)} \in \{0, 1\}, \quad \forall j \in [M], x, y \in \mathcal{X} \\ & && F_j \geq 0 \quad \forall j \in [M], \\ & && \sum_{j=1}^M F_j = \mathbf{1}_B, \\ & && \sum_{x \in \mathcal{X}} P_{xy}^{(j)} = 1, \quad \forall j \in [M], y \in \mathcal{X} \\ & && \sum_{y \in \mathcal{X}} P_{xy}^{(j)} = 1, \quad \forall j \in [M], x \in \mathcal{X} \end{aligned} \quad (6.90)$$

Note that all the constraints are semidefinite or linear, except that each element  $P_{xy}^{(j)}$  is a binary variable:  $P_{xy}^{(j)} \in \{0, 1\}$ , which is a particularly simple type of discrete constraint. The non-linearity in the objective function, however, persists. To remove this, we take advantage of the fact that the  $P_{xy}^{(j)}$  are binary. In particular, [BDNK19, Equations (22)–(24)] provide a clever trick to turn objective functions with terms of the form  $zx$  where  $z$  is a binary variable and  $x$  a continuous variable into objective functions of a continuous variable  $y$  subject to four affine constraints (in terms of  $x$  and  $z$ ), as long as  $x$  is bounded by known constants. We reproduce this argument in the following.

We first write the objective function entirely in terms of scalar quantities:

$$\sum_{x,y \in \mathcal{X}, j \in [M]} p_X(x) P_{xy}^{(j)} c_y \operatorname{tr}[F_j \rho_B^x] = \sum_{k, \ell \in [d_B]} \sum_{x,y \in \mathcal{X}, j \in [M]} p_X(x) (\rho_B^x)_{k\ell} c_y P_{xy}^{(j)} (F_j)_{\ell k} \quad (6.91)$$

Let  $x = (F_j)_{\ell k}$  and  $z = P_{xy}^{(j)} \in \{0, 1\}$ . Then  $|x| \leq \operatorname{tr}[F_j]/2 \leq d_B/2$ . Then  $x_L := -d_B/2$  and  $x_U := d_B/2$  constitute lower and upper bounds to  $x$ , respectively. Hence, the following four inequalities hold trivially:

$$\begin{aligned} z(x - x_L) &\geq 0, \\ (z - 1)(x - x_U) &\geq 0, \\ z(x - x_U) &\leq 0, \\ (z - 1)(x - x_L) &\leq 0. \end{aligned} \quad (6.92)$$

Now, let  $y = xz$ . Then we have

$$\begin{aligned} y - zx_L &\geq 0, \\ y - zx_U &\geq x - x_U, \\ y - zx_U &\leq 0, \\ y - zx_L &\leq x - x_L. \end{aligned} \quad (6.93)$$

On the other hand, let us remove the constraint  $y = xz$ , and consider  $y$  as another variable. Then if  $z = 0$ , the first equation of (6.93) implies that  $y \geq 0$ , while the third implies  $y \leq 0$ , so  $y = 0$ . On the other hand, if  $z = 1$ , then the second equation of (6.93) implies that  $y \geq x$  while the fourth implies that  $y \leq x$ . Hence, either way,  $y = xz$ . Thus, (6.93) is equivalent to  $y = xz$ .

With this transformation, (6.90) can be reformulated as the following.

$$\begin{aligned}
& \text{minimize} && \sum_{k, \ell \in [d_B]} \sum_{x, y \in \mathcal{X}, j \in [M]} p_X(x) (\rho_B^x)_{k\ell} c_{y,j} y_{xy\ell k j} \\
& \text{subject to} && F_j \in \mathcal{M}_{d_B} \quad \forall j \in [M], \\
& && y_{xy\ell k j} \in \mathbb{R}, \quad \forall x, y \in \mathcal{X}, \ell, k \in [d_B], j \in [M], \\
& && P_{xy}^{(j)} \in \{0, 1\}, \quad \forall j \in [M], x, y \in \mathcal{X} \\
& && F_j \geq 0 \quad \forall j \in [M], \\
& && \sum_{j=1}^M F_j = \mathbf{1}_B, \\
& && \sum_{x \in \mathcal{X}} P_{xy}^{(j)} = 1, \quad \forall j \in [M], y \in \mathcal{X} \\
& && \sum_{y \in \mathcal{X}} P_{xy}^{(j)} = 1, \quad \forall j \in [M], x \in \mathcal{X} \\
& && y_{xy\ell k j} + P_{xy}^{(j)} \frac{d_B}{2} \geq 0 \quad \forall x, y \in \mathcal{X}, \ell, k \in [d_B], j \in [M], \\
& && y_{xy\ell k j} - P_{xy}^{(j)} \frac{d_B}{2} \geq (F_j)_{\ell k} - \frac{d_B}{2} \quad \forall x, y \in \mathcal{X}, \ell, k \in [d_B], j \in [M], \\
& && y_{xy\ell k j} - P_{xy}^{(j)} \frac{d_B}{2} \leq 0 \quad \forall x, y \in \mathcal{X}, \ell, k \in [d_B], j \in [M], \\
& && y_{xy\ell k j} + P_{xy}^{(j)} \frac{d_B}{2} \leq (F_j)_{\ell k} + \frac{d_B}{2} \quad \forall x, y \in \mathcal{X}, \ell, k \in [d_B], j \in [M].
\end{aligned} \tag{6.94}$$

This is a mixed-integer SDP, with a number of constraints and variables that is polynomial in  $M, d_B, |\mathcal{X}|$ . Moreover, if  $M \geq d_B^2$ , then as follows from the remark below (6.58), the mixed-integer SDP (6.94) obtains the same optimal value as (6.58), namely  $G_{\bar{c}}(X|B)_\rho$ . Note, however, that mixed-integer SDPs are not in general efficiently solvable; they encompass mixed integer linear programs, which are NP-hard. However, in practice they can sometimes be quickly solved. Since the original SDP formulation (6.58) involves an exponential (in  $|\mathcal{X}|$ ) number of variables (or an exponential number of constraints in its dual formulation (6.85)), (6.94) which instead has a polynomial (in  $|\mathcal{X}|$ ) number of variables may provide a more practical approach in some cases. Mixed-integer SDPs can be solved in various ways; in the code [Han20b], the problem (6.94) is solved using the library Pajarito.jl [CLV18], which proceeds by solving an alternating sequence of mixed-integer linear problems and SDPs.



### The ellipsoid algorithm

The ellipsoid algorithm (see e.g. [GLS93]) provides a theoretical proof that under a strict feasibility assumption, semi-definite programs can be solved in time that scales as a polynomial in: the number of scalar variables and constraints, the log of a 2-norm bound on the feasible points,  $\log(1/\varepsilon)$  where  $\varepsilon$  is the solution tolerance, and the maximum bit length of the scalar entries of the objective and constraints (see e.g. [Wat09, Theorem 4]).

In fact, the ellipsoid algorithm applies quite generally to the optimization of a linear objective function over a convex *feasible region* (which could be described by a domain and constraint functions). The ellipsoid algorithm only requires a *separation oracle* for the feasible region, a subroutine which either asserts that a given point lies within the feasible region, or provides a separating hyperplane between the given point and the feasible region. When the separation oracle can be evaluated in polynomial time, the overall ellipsoid algorithm runs in polynomial time as well (see [GLS93, Corollary 4.2.7]).

In the case of a single positive semi-definite constraint, e.g.  $Y \geq 0$ , a simple separation oracle is given by computing the eigendecomposition of  $Y$  and checking if all of its eigenvalues are non-negative. If so, it returns that  $Y$  is indeed feasible, and otherwise returns the matrix  $C := U \text{diag}(f(\lambda_1), \dots, f(\lambda_d)) U^\dagger$  where  $Y = U \text{diag}(\lambda_1, \dots, \lambda_d) U^\dagger$  is the eigendecomposition of  $Y$ ,  $U$  is unitary,  $\lambda_1, \dots, \lambda_d$  are the eigenvalues, and  $f(x) = 1$  if  $x < 0$  and  $f(x) = 0$  otherwise. The matrix  $C$  has the properties that  $C \geq 0$ ,  $\|C\|_\infty = 1$ , and  $\text{tr}[C^\dagger Y] = \sum_{i=1}^d f(\lambda_i) \lambda_i = \sum_{i: \lambda_i < 0} \lambda_i < 0$ .

In the case of the dual problem (6.85) with  $K = |\mathcal{X}|$ , we have  $|\mathcal{X}|!$  positive semidefinite constraints. Thus, we cannot check all of them together in polynomial time.

The feasibility problem  $Y \leq R_\pi$  for each  $\pi \in S_{|\mathcal{X}|}$  can be written as the following mixed-integer non-linear problem,

$$\begin{aligned} \eta := \text{minimize} \quad & \langle \psi, \left( \sum_{x \in \mathcal{X}} (Pc)_x p_X(x) \rho_B^x - Y \right) \psi \rangle \\ \text{subject to} \quad & \sum_{i \in \mathcal{X}} P_{ij} = \sum_{j \in \mathcal{X}} P_{ij} = 1 \\ & P_{ij} \in \{0\}, \quad i, j \in \mathcal{X} \\ & \psi \in \mathbb{C}^{d_B} \end{aligned}$$

where  $\eta \geq 0$  if and only if  $Y \leq R_\pi$  for all  $\pi \in S_{|\mathcal{X}|}$ , using that a matrix  $M$  satisfies  $M \geq 0$  if and only if  $\langle \psi, M \psi \rangle \geq 0$  for all  $\psi \in \mathbb{C}^{d_B}$ . Since the convex hull of the permutation matrices is given by the doubly stochastic matrices, the discrete constraints can be

relaxed, yielding the following reformulation

$$\begin{aligned}
\eta = \text{minimize} \quad & \langle \psi, \left( \sum_{x \in \mathcal{X}} (Dc)_x p_X(x) \rho_B^x - Y \right) \psi \rangle \\
\text{subject to} \quad & \sum_{i \in \mathcal{X}} D_{ij} = \sum_{j \in \mathcal{X}} D_{ij} = 1 \\
& D_{ij} \geq 0, \quad i, j \in \mathcal{X} \\
& \psi \in \mathbb{C}^{d_B}.
\end{aligned} \tag{6.95}$$

If  $\eta \geq 0$ , then  $Y$  is feasible. Otherwise, the optimal value  $D^*$  can be decomposed as a convex combination of permutations,  $D^* = \sum_i \alpha_i P_i$ , and we must have  $\sum_x (P_i c)_x p_X(x) \rho_B^x \not\geq Y$  for some  $P_i$ , using that the objective is an affine function of  $D$ . The problem (6.95) can be solved by global non-linear optimization solvers such as EAGO.jl [WS17] or SCIP [GABC+20], but not in general in polynomial time.

At each iteration of the ellipsoid algorithm, one must evaluate the separation oracle for some Hermitian matrix  $Y$ . In order to avoid solving (6.95), one may attempt to prove the feasibility or infeasibility of a point  $Y$  by other means. For example, one may search over permutations  $\pi$  heuristically, in order to find  $R_\pi$  such that  $Y \not\geq R_\pi$ . If such a permutation can be identified, then  $Y$  is not feasible, and the problem (6.95) does not need to be solved. Likewise, if one can show that for some  $k \in (1, \dots, |\mathcal{X}|)$ ,

$$Y \leq \sum_{i=1}^k c_{|\mathcal{X}|-i} p_{x_i} \rho_B^{x_i} + c_1 \sum_{x \in \mathcal{X} \setminus \{x_{i_1}, \dots, x_{i_k}\}} p_x \rho_B^x \forall (x_1, \dots, x_k) \in \mathcal{X}_{\neq}^k \tag{6.96}$$

then  $Y$  must be feasible, and again (6.95) does not need to be solved. The number of comparisons required scales as  $|\mathcal{X}|^k$ ; for small choices of  $k$ , this provides an efficient check for feasibility (which may, however, be inconclusive).

## Numerical experiments

We compare numeric implementations of several of the above algorithms on a set of 12 test problems. The code for these experiments can be found at [Han20b]. Each problem has  $p \equiv u$ , the uniform distribution  $u := (1/|\mathcal{X}|, \dots, 1/|\mathcal{X}|)$ , for simplicity. The states are chosen as

1. Two random qubit density matrices
2. Two random qutrit density matrices

3. Three pure qubits chosen equidistant within one plane of the Bloch sphere (the “Y-states”), i.e.

$$\cos\left(j\frac{2\pi}{3}\right)|0\rangle + \sin\left(j\frac{2\pi}{3}\right)|1\rangle, \quad j = 1, 2, 3$$

4. Three random qubit density matrices
5. Three random qutrit density matrices
6. The four BB84 states,  $|0\rangle$ ,  $|1\rangle$ , and  $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$

as well as the “tensor-2” case of

$$\{\rho \otimes \sigma : \rho, \sigma \in S\} \tag{6.97}$$

for each of the six sets  $S$  listed above, corresponding to the guesswork problem with quantum side information associated to  $\rho_{XB}^{\otimes 2}$ , where  $\rho_{XB}$  is the state associated to the original guesswork problem with quantum side information. The random states were chosen uniformly at random (i.e. according to the Haar measure).

The exponentially-large SDP formulation (and its dual), the mixed-integer SDP algorithm, and the active set method were compared, with several choices of parameters and underlying solvers. The mixed-integer SDP formulation was evaluated with  $M = d_B$  (yielding an upper bound),  $M = d_B^2$  (yielding the optimal value), with the Pajarito mixed-integer SDP solver [CLV18], using Convex.jl (version 0.12.7) [UMZH+14] to formulate the problem. Pajarito proceeds by solving mixed-integer linear problems (MILP) and SDPs as subproblems, and thus uses both a MILP solver and an SDP solver as subcomponents. Pajarito provides two algorithms: an iterative algorithm, which alternates between solving MILP and SDP subproblems, and solving a single branch-and-cut problem in which SDP subproblems are solved via so-called lazy callbacks to add cuts to the mixed-integer problem. The latter is called “mixed-solver drives” (MSD) in the Pajarito documentation. We tested three configurations of Pajarito (version 0.7.0):

- (c1) Gurobi (version 9.0.3) as the MILP solver and MOSEK (version 8.1.0.82) as the SDP solver, with Pajarito’s MSD algorithm
- (c2) Gurobi as the MILP solver and MOSEK as the SDP solver, with Pajarito’s iterative algorithm, with a relative optimality gap tolerance of 0,
- (o) Cbc [jVRS+19] (version 2.10.3) as the MILP solver, and SCS [OCPB16] (version 2.1.1) as the SDP solver, with Pajarito’s iterative algorithm

Here, ‘c’ stands for commercial, and ‘o’ for open-source. In the configuration (c1), Gurobi was set to have with a relative optimality gap tolerance of  $10^{-5}$  and in (c2), a relative optimality gap tolerance of 0. In both configurations, Gurobi was given an absolute linear-constraint-wise feasibility tolerance of  $10^{-8}$ , and an integrality tolerance of  $10^{-9}$ . These choices of parameters match those made in [CLV18]. Cbc was given an integrality tolerance of  $10^{-8}$ , and SCS’s (normalized) primal, dual residual and relative gap were set to  $10^{-6}$  for each problem. The default parameters were used otherwise. Note the MSD option was not used with Cbc, since the solver does not support lazy callbacks.

For the (exponentially large) SDP primal and dual formulations, the problems were solved with both MOSEK and SCS, and likewise with the active-set upper bound.

The active set method uses simulated annealing to iteratively add violated constraints to the problem to find an upper bound, as described in Section 5.5.1, and uses a maximum-time parameter  $t_{\max}$  to stop iterating when the estimated time of finding another constraint to add would cause the running time of to exceed the maximum-time<sup>3</sup>. This provides a way to compare the improvement (or lack thereof) of running the algorithm for more iterations. The algorithm also terminates when a violated constraint cannot be found after 50 runs of simulated annealing (started each time with different random initial conditions). Here, the problems were solved with three choices of  $t_{\max}$ , 20 s, 60 s, and 240 s.

The exact answer was not known analytically for most of these problems, so the average relative error was calculated by comparing to the mean of the solutions (excluding the active-set method and the MISDP with  $M = d_B$ , which only give an upper bound in general). In the BB84 case, in which the solution is known exactly (see [HKDW20]), the solutions obtained here match the analytic value to a relative tolerance of at least  $10^{-7}$ .

The problems were run sequentially on a 4-core desktop computer (Intel i7-6700K 4.00GHz CPU, with 16 GB of RAM, on Ubuntu-20.04 via Windows Subsystem for Linux, version 2), via the programming language Julia [BEKS17] (version 1.5.1), with a 5 minute time limit. The results are summarized in Table 6.1, and presented in more detail in Table 6.2 and Table 6.3.

One can see that the MISDP problems were harder solve than the corresponding SDPs for these relatively small problem instances. The MISDPs have the advantage of finding extremal solutions, however, in the case  $M = d_B$ , and may scale better to large instances. Additionally, the active-set upper bound performed fairly well, finding

<sup>3</sup>The maximum time can still be exceeded, since at least one iteration must be performed and the estimate can be wrong.

feasible points within 20 % of the optimum in all cases, with only  $t_{\max} = 20$  s, and often finding near-optimal solutions. It was also the only method able to scale to the largest instances tested, such as two copies of the BB84 states (which involves 16 quantum states in dimension 4, and which the SDP formulation has  $16!$  variables.). In general, the commercial solvers performed better than the open source solvers, with the notable exception of the active-set upper bound with MOSEK, in which 2 more problems timed out than with SCS. This could be due to SCS being a first-order solver which can therefore possibly scale to larger problem instances than MOSEK, which is a second-order solver.

### 6.3.6 Entropic bounds

In this section, we use the results of Section 6.3.2 to obtain one-shot and asymptotic entropic bounds on the guesswork with side information in terms of measured versions of bounds known in the classical case.

#### One-shot bounds

Arikan [Ari96] showed that

$$\frac{1}{1 + \ln |\mathcal{X}|} \exp(H_{\frac{1}{2}}^{\uparrow}(X|Y)_p) \leq G(X|Y)_p \leq \exp(H_{\frac{1}{2}}^{\uparrow}(X|Y)_p) \quad (6.98)$$

where  $H_{\alpha}^{\uparrow}(X|Y)_p$  for  $\alpha \in (0, 1) \cup (1, \infty)$  denotes the following  $\alpha$ -conditional entropy of a joint distribution  $p_{XY}$  given by

$$H_{\alpha}^{\uparrow}(X|Y) = \frac{\alpha}{1 - \alpha} \ln \left( \sum_{y \in \mathcal{Y}} \left( \sum_{x \in \mathcal{X}} p_{XY}(x, y)^{\alpha} \right)^{1/\alpha} \right) = \sup_{q_Y} [-D_{\alpha}(p_{XY} \| \mathbf{1}_X \otimes q_Y)] \quad (6.99)$$

where the supremum is over probability distributions  $q_Y$  on  $\mathcal{Y}$ , and  $D_{\alpha}$  is the  $\alpha$ -Rényi relative entropy,

$$D_{\alpha}(p_X \| q_X) = \frac{1}{\alpha - 1} \ln \left( \sum_x p_X(x)^{\alpha} q_X(x)^{1-\alpha} \right). \quad (6.100)$$

The second equality of (6.99) follows from of [FB14, Theorem 4].

Algorithm	Parameters	average rel. error	average time	number solved	number timed out	number errored out
MISDP ( $d_B$ )	Pajarito (c1)	0 %	23.74 s	6	6	0
	Pajarito (c2)	0 %	24.03 s	6	6	0
	Pajarito (o)	0 %	45.05 s	6	6	0
MISDP ( $d_B^2$ )	Pajarito (c1)	0 %	35.27 s	5	7	0
	Pajarito (c2)	0 %	27.97 s	4	8	0
	Pajarito (o)	0 %	131.35 s	4	8	0
SDP	MOSEK	0 %	8.99 s	8	3	1
	SCS	0 %	9.08 s	8	3	1
SDP (dual)	MOSEK	0 %	8.74 s	8	3	1
	SCS	0 %	8.59 s	8	3	1
Active set	$t_{\max} = 20$ s	6.80 %	16.08 s	10	2	0
upper bound	$t_{\max} = 60$ s	6.79 %	19.03 s	10	2	0
(MOSEK)	$t_{\max} = 240$ s	6.80 %	26.17 s	10	2	0
Active set	$t_{\max} = 20$ s	6.09 %	33.24 s	12	0	0
upper bound	$t_{\max} = 60$ s	6.09 %	35.78 s	12	0	0
(SCS)	$t_{\max} = 240$ s	6.09 %	34.30 s	11	1	0

Table 6.1 Comparison of average relative error and average solve time for the 12 problems discussed above. A problem is considered “timed out” if an answer is not obtained in 5 minutes, and “errored out” if the solution was not obtained due to errors (such as running out of RAM). The average relative error, which was rounded to two decimal digits, and the time taken are calculated only over the problems which were solved by the given algorithm and choice of parameters. “MISDP ( $d_B$ )” refers to the choice  $M = d_B$ , and likewise “MISDP ( $d_B^2$ )” refers to the choice  $M = d_B^2$ .

Algorithm	Parameters	Two random qubits	Two random qutrits	Y-states
MISDP ( $d_B$ )	Pajarito (c1)	23.63 s, timeout	23.60 s, timeout	23.56 s, timeout
	Pajarito (c2)	22.99 s, timeout	23.31 s, timeout	23.21 s, timeout
	Pajarito (o)	23.47 s, timeout	24.77 s, timeout	26.15 s, timeout
MISDP ( $d_B^2$ )	Pajarito (c1)	24.49 s (0.00 %), timeout	31.40 s (0.00 %), timeout	26.27 s (0.00 %), timeout
	Pajarito (c2)	25.02 s (0.00 %), timeout	31.39 s (0.00 %), timeout	29.08 s (0.00 %), timeout
	Pajarito (o)	26.54 s (0.00 %), timeout	212.79 s (0.00 %), timeout	141.14 s (0.00 %), timeout
SDP	MOSEK	8.69 s, 8.84 s	8.78 s, 9.23 s	9.33 s, timeout
	SCS	9.00 s, 8.90 s	8.44 s, 11.22 s	8.98 s, timeout
SDP (dual)	MOSEK	8.46 s, 8.63 s	8.54 s, 8.83 s	9.11 s, timeout
	SCS	8.76 s, 8.32 s	8.33 s, 9.20 s	8.74 s, timeout
Active set upper bound (MOSEK)	$t_{\max} = 20$ s	8.76 s (19.5 %), 10.41 s (1.5 %)	8.89 s (19.5 %), timeout	9.72 s (0 %), 34.25 s (? %)
	$t_{\max} = 60$ s	10.91 s (19.5 %), 10.41 s (1.9 %)	8.87 s (19.5 %), timeout	9.66 s (0 %), 31.00 s (? %)
	$t_{\max} = 240$ s	9.47 s (19.5 %), 10.40 s (1.5 %)	8.90 s (19.5 %), timeout	9.81 s (0 %), 30.26 s (? %)
Active set upper bound (SCS)	$t_{\max} = 20$ s	9.04 s (19.5 %), 10.92 s (1.5 %)	8.70 s (19.5 %), 101.06 s (1.1 %)	9.23 s (0 %), 82.84 s (? %)
	$t_{\max} = 60$ s	9.07 s (19.5 %), 10.22 s (1.9 %)	8.66 s (19.5 %), 32.94 s (1.2 %)	9.18 s (0 %), 50.79 s (? %)
	$t_{\max} = 240$ s	9.04 s (19.5 %), 10.02 s (1.9 %)	8.79 s (19.5 %), 22.69 s (1.2 %)	9.31 s (0 %), 37.36 s (? %)

Table 6.2 The individual timings for each algorithm and choice of settings on problems (1)–(3), and the corresponding “tensor-2” problems discussed at (6.97). For each algorithm, the running time of the original problem is given followed by the running time on the “tensor-2” problem, e.g. the SDP formulation with MOSEK on the two random qubits problem was solved in 8.69 seconds, and in 8.84 seconds for the corresponding tensor-2 problem. “timeout” is written whenever the problem was not solved within 5 minutes. For the active set algorithms, the relative error is also given for each problem in parenthesis. Note that the MISDP formulation with  $M = d_B$  is also only known to be an upper bound, but a relative error of less than  $10^{-5}$  in each instance, so the relative errors are omitted. Lastly, the relative error is written as ? % in the case that only an upper bound was obtained.

Algorithm	Parameters	Three random qubits	Three random qutrits	BB84 states
MISDP ( $d_B$ )	Pajarito (c1)	23.63 s, timeout	24.49 s, timeout	23.51 s, timeout
	Pajarito (c2)	23.17 s, timeout	26.50 s, timeout	25.01 s, timeout
	Pajarito (o)	27.11 s, timeout	95.01 s, timeout	73.80 s, timeout
MISDP ( $d_B^2$ )	Pajarito (c1)	25.82 s (0.00 %), timeout	timeout, timeout	68.35 s (0.00 %), timeout
	Pajarito (c2)	26.38 s (0.00 %), timeout	timeout, timeout	timeout, timeout
	Pajarito (o)	144.93 s (0.00 %), timeout	timeout, timeout	timeout, timeout
SDP	MOSEK	9.35 s, timeout	8.82 s, timeout	8.87 s, error
	SCS	9.09 s, timeout	8.46 s, timeout	8.54 s, error
SDP (dual)	MOSEK	9.14 s, timeout	8.55 s, timeout	8.62 s, error
	SCS	8.86 s, timeout	8.25 s, timeout	8.23 s, error
Active set upper bound (MOSEK)	$t_{\max} = 20$ s	9.73 s (1.3 %), 32.29 s (? %)	9.50 s (5.8 %), timeout	9.51 s (0 %), 27.72 s (? %)
	$t_{\max} = 60$ s	9.69 s (1.3 %), 24.12 s (? %)	9.45 s (5.8 %), timeout	9.77 s (0 %), 66.47 s (? %)
	$t_{\max} = 240$ s	9.69 s (1.3 %), 30.42 s (? %)	9.45 s (5.8 %), timeout	9.49 s (0 %), 133.81 s (? %)
Active set upper bound (SCS)	$t_{\max} = 20$ s	9.44 s (1.3 %), 35.66 s (? %)	9.67 s (5.8 %), 84.43 s (? %)	9.09 s (0 %), 28.76 s (? %)
	$t_{\max} = 60$ s	9.44 s (1.3 %), 54.60 s (? %)	9.11 s (5.8 %), 155.51 s (? %)	9.29 s (0 %), 70.57 s (? %)
	$t_{\max} = 240$ s	8.89 s (1.3 %), 75.43 s (? %)	9.01 s (5.8 %), timeout	9.15 s (0 %), 177.64 s (? %)

Table 6.3 The individual timings for each algorithm and choice of settings on problems (4)–(6). See Table 6.2 for a description of the quantities shown. Here, “error” means the solution was not obtained due to an error (such as running out of memory).



Arikan's bound (6.98) applies to each  $G(X|Y)_p$  in (6.48), and hence by minimizing over the POVMs  $\{E_y\}_{y \in \mathcal{Y}}$ , we obtain

$$\frac{1}{1 + \ln |\mathcal{X}|} \exp(H_{\frac{1}{2}}^{\uparrow, M}(X|B)_\rho) \leq G(X|B)_\rho \leq \exp(H_{\frac{1}{2}}^{\uparrow, M}(X|B)_\rho), \quad (6.101)$$

where for  $\alpha \in (0, 1) \cup (1, \infty)$ ,  $H_\alpha^{\uparrow, M}(X|B)_\rho$  is the  $B$ -measured conditional  $\alpha$ -Rényi entropy, defined by

$$H_\alpha^{\uparrow, M}(X|B)_\rho := \inf_{\{E_y\}_{y \in \mathcal{Y}}} H_\alpha^\uparrow(X|Y)_p, \quad (6.102)$$

where the infimum is taken over POVMs  $\{E_y\}_{y \in \mathcal{Y}}$  and  $p_{XY}(x, y) = p_X(x) \operatorname{tr}[E_y \rho_B^x]$  is the joint probability distribution obtained by measuring the  $B$  part of  $\rho_{XB}$  via  $\{E_y\}_{y \in \mathcal{Y}}$ .

*Remark.* We may expand this quantity as

$$H_\alpha^{\uparrow, M}(X|B)_\rho = \inf_{\{E_y\}_{y \in \mathcal{Y}}} \sup_{q_Y} [-D_\alpha(p_{XY} \| \mathbf{1}_X \otimes q_Y)], \quad (6.103)$$

where  $p_{XY}$  is induced by the measurement of  $\rho_{XB}$ . This quantity seems to be different from the *conditional entropy induced by the measured Rényi divergence*, namely

$$H_{D_\alpha^M}^\uparrow(X|B)_\rho := \sup_{\sigma_B} -D_\alpha^M(\rho_{XB} \| \mathbf{1}_X \otimes \sigma_B), \quad (6.104)$$

where the supremum is over states on the  $B$  system, and for any pair of states  $(\rho, \sigma)$ ,

$$D_\alpha^M(\rho \| \sigma) := \sup_{\{E_z\}_z} D_\alpha(\{\operatorname{tr}[E_z \rho]\}_z \| \{\operatorname{tr}[E_z \sigma]\}_z) \quad (6.105)$$

is the *measured  $\alpha$ -Rényi divergence*. Indeed, the latter quantity may be expanded to obtain

$$H_{D_\alpha^M}^\uparrow(X|B)_\rho = \sup_{\sigma_B} \inf_{\{E_z\}_z} [-D_\alpha(\{\operatorname{tr}[E_z \rho_{XB}]\}_z \| \{\operatorname{tr}[E_z \mathbf{1}_X \otimes \sigma_B]\}_z)]. \quad (6.106)$$

From the min-max inequality, and the fact that collective measurements on  $XB$  can simulate measurements on  $B$  alone, we have

$$H_{D_\alpha^M}^\uparrow(X|B)_\rho \leq H_\alpha^{\uparrow, M}(X|B)_\rho. \quad (6.107)$$

### Asymptotic analysis

We can consider the asymptotic setting in which a Bob receives a sequence of product states  $\rho_{\vec{B}} := \rho_B^{x_1} \otimes \cdots \otimes \rho_B^{x_n}$ , with probability  $p_X(x_1) \cdots p_X(x_n)$  and aims to guess the full sequence  $\vec{x} = (x_1, \dots, x_n)$ . In this case, the problem is characterized by the c-q state  $\rho_{XB}^{\otimes n}$ . The 1-shot bounds (6.101) give us

$$-\frac{1}{n} \ln(1 + n \ln(|\mathcal{X}|)) + \frac{1}{n} H_{\frac{1}{2}}^{\uparrow, M}(X^n | B^n)_{\rho^{\otimes n}} \leq \frac{1}{n} \ln G(X^n | B^n)_{\rho^{\otimes n}} \leq \frac{1}{n} H_{\frac{1}{2}}^{\uparrow, M}(X^n | B^n)_{\rho^{\otimes n}} \quad (6.108)$$

where  $H_{\frac{1}{2}}^{\uparrow, M}(X^n | B^n)_{\rho^{\otimes n}}$  can involve collective measurements on the system  $B^n$ . Taking  $n \rightarrow \infty$ , we obtain

$$\lim_{n \rightarrow \infty} \frac{1}{n} \ln G(X^n | B^n)_{\rho^{\otimes n}} = \lim_{n \rightarrow \infty} \frac{1}{n} H_{\frac{1}{2}}^{\uparrow, M}(X^n | B^n)_{\rho^{\otimes n}}, \quad (6.109)$$

assuming that the limit on the right-hand side exists.

Note that we can bound

$$\frac{1}{n} H_{\frac{1}{2}}^{\uparrow, M}(X^n | B^n)_{\rho^{\otimes n}} \leq \frac{1}{n} \inf_{\{E_y\}_{y \in \mathcal{Y}}} H_{\frac{1}{2}}^{\uparrow}(X^n | Y^n)_{p^{\otimes n}} \quad (6.110)$$

$$= \inf_{\{E_y\}_{y \in \mathcal{Y}}} H_{\frac{1}{2}}^{\uparrow}(X | Y)_p \quad (6.111)$$

$$= H_{\frac{1}{2}}^{\uparrow, M}(X | B)_{\rho} \quad (6.112)$$

where the first inequality follows from the fact that product measurements are a special case of collective measurements, and the first equality follows from the additivity of the classical Rényi entropy (Proposition 1 of [Ari96]), and the third by the definition of  $H_{\frac{1}{2}}^{\uparrow, M}(X | B)_{\rho}$ . Moreover, by the data-processing inequality [FL13],

$$\frac{1}{n} H_{\frac{1}{2}}^{\uparrow, M}(X^n | B^n)_{\rho^{\otimes n}} \geq \frac{1}{n} \tilde{H}_{\frac{1}{2}}^{\uparrow}(X^n | B^n)_{\rho^{\otimes n}} = \tilde{H}_{\frac{1}{2}}^{\uparrow}(X | B)_{\rho}, \quad (6.113)$$

where the conditional Rényi entropy  $\tilde{H}_{\alpha}^{\uparrow}(C | D)_{\sigma}$  of a bipartite state  $\sigma_{CD}$  is defined as

$$\tilde{H}_{\alpha}^{\uparrow}(C | D)_{\sigma} = \sup_{\omega_D} \left[ -\tilde{D}_{\alpha}(\sigma_{CD} \| \mathbf{1}_C \otimes \omega_D) \right], \quad (6.114)$$

with the optimization with respect to states  $\omega_D$  and the sandwiched Rényi relative entropy defined as [MDSF+13; WWY14]:

$$\tilde{D}_{\alpha}(X \| Y) = \frac{1}{\alpha - 1} \ln \text{tr}[(Y^{(1-2\alpha)/\alpha} X Y^{(1-2\alpha)/\alpha})^{\alpha}]. \quad (6.115)$$

The equality in (6.113) follows from the additivity of  $\widetilde{H}_{\frac{1}{2}}^\uparrow$  under tensor products (see, e.g., Corollary 5.2 of [Tom16]). Hence, we obtain

$$\widetilde{H}_{\frac{1}{2}}^\uparrow(X|B)_\rho \leq \lim_{n \rightarrow \infty} \frac{1}{n} \ln G(X^n|B^n)_{\rho^{\otimes n}} \leq H_{\frac{1}{2}}^{\uparrow, M}(X|B)_\rho. \quad (6.116)$$

In the classical case, (6.7), both the left and right-hand sides reduce to

$$H_{\frac{1}{2}}^\uparrow(X|Y)_p \quad (6.117)$$

where  $p$  is the underlying classical distribution of (6.7). Hence, these bounds recover Proposition 5 of [Ari96].



# Chapter 7

## General concave functions maximized over the $\varepsilon$ -ball

The main results of sections Chapters 3 and 5, are essentially fully classical in that the first step of the proofs in the case of quantum states simply consists of a reduction to the case of probability vectors in  $\mathbb{R}^d$ . In this section, we prove Theorem 3.4.1 which was used to motivate the construction of  $\rho_\varepsilon^*$  by finding necessary and sufficient conditions for maximizing concave functions over the  $\varepsilon$ -ball. Moreover, we can treat—to some extent—functions which are not unitarily invariant and depend on the eigenprojections as well as the eigenvalues of a state. In particular, in Corollary 7.3.2 we obtain necessary and sufficient conditions for a state  $\xi$  to maximize the conditional entropy over the  $\varepsilon$ -ball. Unfortunately, we will not be able to determine a formula for a specific maximizer, and we do not expect such a maximizer to have the universality of  $\rho_\varepsilon^*$  (which maximizes any Schur concave function over the  $\varepsilon$ -ball  $B_\varepsilon(\rho)$ ). A similar approach of investigating the critical points of entropic quantities is undertaken in [Teh20].

### 7.1 Tools from convex optimization

Let  $\mathcal{X}$  be a real Hilbert space with inner product  $\langle \cdot, \cdot \rangle$  and  $f : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$  be a function. In our applications, we take  $\mathcal{X} = \mathcal{B}_{\text{sa}}(\mathcal{H})$  equipped with the Hilbert-Schmidt inner product  $\langle \cdot, \cdot \rangle_{\text{HS}}$ . Let  $\text{dom } f = \{x \in \mathcal{X} : f(x) < \infty\}$  and assume  $\text{dom } f \neq \emptyset$ . Let  $\mathcal{X}^* = \mathcal{B}(\mathcal{X}, \mathbb{R})$  the set of bounded linear maps from  $\mathcal{X}$  to  $\mathbb{R}$ , equipped with the dual norm  $\|\ell\|_* = \sup_{\|x\|=1} |\ell(x)|$  for  $\ell \in \mathcal{X}^*$ . Since  $\mathcal{X}$  is a Hilbert space, by the Riesz-Fréchet representation, for each  $\ell \in \mathcal{X}^*$  there exists a unique  $u_\ell \in \mathcal{X}$  such that  $\ell(x) = \langle u_\ell, x \rangle$  for all  $x \in \mathcal{X}$ . We call  $u_\ell$  the dual vector for  $\ell$  (in particular the Hilbert-Schmidt dual in the case of  $\mathcal{X} = (\mathcal{B}_{\text{sa}}(\mathcal{H}), \langle \cdot, \cdot \rangle_{\text{HS}})$ ). We say  $f$  is *lower semicontinuous* if  $\liminf_{x \rightarrow x_0} f(x) \geq f(x_0)$ .

for each  $x_0 \in \mathcal{X}$ . The *directional derivative* of  $f$  at  $x \in \text{dom } f$  in the direction  $h \in \mathcal{X}$  is given by

$$f'(x; h) = \lim_{t \downarrow 0} \frac{1}{t} [f(x + th) - f(x)]. \quad (7.1)$$

If  $f$  is convex, this limit exists in  $\mathbb{R} \cup \{\pm\infty\}$ . If the map  $\phi_x(h) := \mathcal{X} \ni h \mapsto f'(x; h)$  is linear and continuous, then  $f$  is called *Gâteaux-differentiable* at  $x \in \mathcal{X}$ . Moreover, if  $f$  is Gâteaux differentiable at every  $x \in A \subseteq \mathcal{X}$ , then  $f$  is said to be Gâteaux-differentiable on  $A$ . If  $f$  is convex and continuous at  $x$ , it can be shown that the map  $\phi_x$  is finite and continuous. However, a continuous and convex function may not be Gâteaux-differentiable. For example,  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) = |x|$  has  $\phi_0(h) = \lim_{t \downarrow 0} \frac{1}{t} [f(x + th) - f(x)] = \lim_{t \downarrow 0} \frac{1}{t} |th| = |h|$  which is nonlinear. If  $f$  is Gâteaux-differentiable at  $x$ , we call the dual to  $\phi_x$  as the *Gâteaux gradient* of  $f$  at  $x$ , written  $\nabla f(x) \in \mathcal{X}$ . That is,  $\langle \nabla f(x), h \rangle = \phi_x(h)$  for each  $h \in \mathcal{X}$ .

The function  $f$  is Fréchet-differentiable at  $x$  if there is  $y \in \mathcal{X}$  such that

$$\lim_{\|r\| \rightarrow 0} \frac{1}{\|r\|} |f(x + r) - f(x) - \langle y, r \rangle| = 0 \quad (7.2)$$

and in this case, one writes  $Df(x) \in \mathcal{X}^*$  for the map  $\mathcal{X} \ni z \mapsto Df(x)z = \langle y, z \rangle$ . If  $f$  is Fréchet-differentiable at  $x$ , then by taking  $r = th$  in (7.2) we find  $\langle y, h \rangle = f'(x, h)$  and therefore  $f$  is Gâteaux-differentiable at  $x$  with  $\langle \nabla f(x), h \rangle = Df(x)h$ .

By regarding differentiability as the existence of a linear approximation at a point, we can generalize it by defining a notion of a linear subapproximation at a point.

**Definition 7.1.1.** The *subgradient* of a function  $f : \mathcal{X} \rightarrow \mathbb{R}$  at  $x$  is the set

$$\partial f(x) = \{u \in \mathcal{X} : f(y) - f(x) \geq \langle u, y - x \rangle \ \forall y \in \mathcal{X}\} \subseteq \mathcal{X}. \quad (7.3)$$

For a convex function  $f$ , the following properties hold:

- if  $f$  is continuous at  $x$ , then  $\partial f(x)$  is bounded and nonempty.
- if  $f$  is Gâteaux-differentiable at  $x$ , then  $\partial f(x) = \{\nabla f(x)\}$ .

We briefly prove the second point here. Assume  $f$  is Gâteaux-differentiable at  $x$ . Then

$$\langle \nabla f(x), y - x \rangle = f'(x, y - x) = \lim_{t \downarrow 0} \frac{1}{t} [f((1 - t)x + ty) - f(x)].$$

By convexity,  $f((1 - t)x + ty) \leq (1 - t)f(x) + tf(y)$ . Therefore,  $\langle \nabla f(x), y - x \rangle \leq f(y) - f(x)$ ; hence,  $\nabla f(x) \in \partial f(x)$ .

On the other hand, given  $u \in \partial f$ ,  $h \in \mathcal{X}$ , and  $t > 0$ , we can set  $y = th + x$ . Then

$$\langle u, th \rangle = \langle u, y - x \rangle \leq f(y) - f(x) = f(x + th) - f(x).$$

Dividing by  $t$  and taking the limit  $t \downarrow 0$  yields  $u = \nabla f(x)$ .

Fermat's Rule of convex optimization theory, stated below, provides a simple characterization of the minimum of a function in terms of the zeroes of its subgradient. Moreover, since the subgradient of a Gâteaux differentiable function consists of a single element, namely, its Gâteaux gradient, finding its minimizer amounts to showing that its Gâteaux gradient is equal to zero.

**Theorem 7.1.2** (Fermat's Rule). *Consider a function  $f : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$  with  $\text{dom } f \neq \emptyset$ . Then  $\hat{x} \in \mathcal{X}$  is a global minimizer of  $f$  if and only if  $0_{\mathcal{X}} \in \partial f(\hat{x})$ , where  $0_{\mathcal{X}}$  is the zero vector of  $\mathcal{X}$ .*

*Proof.*  $0_{\mathcal{X}} \in \partial f(\hat{x})$  if and only if

$$f(y) - f(\hat{x}) \geq \langle 0, y - \hat{x} \rangle = 0$$

for every  $y \in \mathcal{X}$ , i.e. if and only if  $\hat{x}$  is a global minimizer of  $f$ .  $\square$

The following result proves useful in computing the subgradient of a sum of convex functions (see e.g. [Pey15, Theorem 3.30] for the proof).

**Theorem 7.1.3** (Moreau-Rockafellar). *Let  $f, g : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$  be convex and lower semicontinuous, with non-empty domains. For each  $A \in \mathcal{X}$ , we have*

$$\partial f(A) + \partial g(A) := \{a + b : a \in \partial f(A), b \in \partial g(A)\} \subseteq \partial(f + g)(A). \quad (7.4)$$

*Equality holds for every  $A \in \mathcal{X}$  if  $f$  is continuous at some  $A_0 \in \text{dom}(g)$ .*

## 7.2 Gâteaux-differentiable functions

Let  $\mathcal{F}$  denote the class of functions  $\varphi : \mathcal{D}(\mathcal{H}) \rightarrow \mathbb{R}$  which are concave and continuous, and Gâteaux-differentiable on  $\mathcal{D}_+(\mathcal{H})$ . Note that these functions need *not* be Schur concave, and in particular, need not be unitarily invariant. These include the following:

- The von Neumann entropy  $\xi \mapsto S(\xi) := -\text{tr}(\xi \log \xi)$ . In Lemma 7.3.7, we show for  $\xi > 0$ ,

$$\nabla S(\xi) = -\log \xi - \frac{1}{\log_e(2)} \mathbf{1}.$$

- The conditional entropy  $\xi_{AB} \mapsto S(A|B)_\xi := S(\xi_{AB}) - S(\xi_B)$ , for which

$$\nabla S(A|B)(\xi_{AB}) = -(\log \xi_{AB} - \mathbf{1}_A \otimes \log \xi_B)$$

as shown in Corollary 7.3.8. Note that for  $\varphi(\xi) := S(A|B)_\xi$ , the conditional entropy satisfies the interesting property that

$$\varphi(\xi) = \langle \nabla \varphi(\xi), \xi \rangle_{\text{HS}}. \quad (7.5)$$

- The  $\alpha$ -Rényi entropy for  $\alpha \in (0, 1)$ . The map  $\xi \mapsto S_\alpha(\xi)$  is concave for  $\alpha \in (0, 1)$ , continuous, and has Gâteaux gradient

$$\nabla S_\alpha(\xi) = \frac{\alpha}{1-\alpha} \frac{1}{\ln 2} \frac{1}{\text{tr}[\xi^\alpha]} \xi^{\alpha-1}$$

by Lemma 7.3.9. Note that the  $\alpha$ -Rényi entropy for  $\alpha > 1$  is not concave.

- The function  $-T_\alpha$ , for  $\xi \mapsto T_\alpha(\xi) := \text{tr}[\xi^\alpha]$  and  $\alpha > 1$ . This map is concave for  $\alpha > 1$ , continuous, and by Lemma 7.3.9, has Gâteaux gradient

$$\nabla(-T_\alpha)(\xi) = -\alpha \xi^{\alpha-1}.$$

A state  $\xi^* \in B_\varepsilon(\rho)$  maximizes  $-T_\alpha$  over  $B_\varepsilon(\rho)$  if and only if  $\xi^*$  minimizes  $T_\alpha$  over  $B_\varepsilon(\rho)$ . For  $\alpha > 1$ , minimizing  $T_\alpha$  is equivalent to maximizing  $S_\alpha(\xi)$ , using that  $x \mapsto \frac{1}{1-\alpha} \log x$  is decreasing. Therefore,  $\xi^*$  maximizes  $-T_\alpha$  over  $B_\varepsilon(\rho)$  if and only if it maximizes  $S_\alpha$  over  $B_\varepsilon(\rho)$ . Thus, by considering the function  $f(\xi) = -T_\alpha(\xi)$  for  $\alpha > 1$  in Theorem 3.4.1, one can establish conditions for  $\xi^*$  to maximize the  $\alpha$ -Rényi entropy for  $\alpha > 1$ .

### 7.3 Optimality conditions for maximizing concave functions over the trace-ball

Theorem 3.4.1 is a condensed form of the following theorem, whose proof we include below.

**Theorem 7.3.1.** *Let  $f : \mathcal{D}(\mathcal{H}) \rightarrow \mathbb{R}$  be a concave, continuous function which is Gâteaux-differentiable on  $\mathcal{D}_+(\mathcal{H})$ . Given a state  $\rho \in \mathcal{D}(\mathcal{H})$  and  $\varepsilon > 0$ , for any  $\xi \in B_\varepsilon^+(\rho)$ , we have*

$$\text{tr}[\nabla f(\xi)(\xi - \rho)] \leq \varepsilon[\lambda_+(\nabla f(\xi)) - \lambda_-(\nabla f(\xi))]. \quad (7.6)$$



Furthermore, the following are equivalent:

1. Equality in (7.6),
2.  $\xi \in \operatorname{argmax}_{B_\varepsilon(\rho)} f$ ,
3. (a) Either  $\frac{1}{2}\|\xi - \rho\|_1 = \varepsilon$  or  $\nabla f(\xi) = \lambda \mathbf{1}$  for some  $\lambda \in \mathbb{R}$ , and  
 (b) we have

$$\pi_\pm \nabla f(\xi) \pi_\pm = \lambda_\pm (\nabla f(\xi)) \pi_\pm$$

where  $\pi_\pm$  is the projection onto the support of  $\Delta_\pm$ , and where  $\Delta = \Delta_+ - \Delta_-$  is the Jordan decomposition of  $\Delta := \xi - \rho$ .

A corollary of this theorem concerns the conditional entropy

$$S(A|B)_\xi := S(AB)_\xi - S(B)_\xi \equiv S(\xi_{AB}) - S(\xi_B), \quad \xi_B := \operatorname{tr}_A \xi_{AB}$$

of a bipartite state  $\xi_{AB}$ . It corresponds to the choice  $\varphi(\xi_{AB}) = S(A|B)_\xi$  whose Gâteaux gradient  $L_\xi := \nabla \varphi(\xi_{AB})$  is given by  $L_{\xi_{AB}} = -(\log \xi_{AB} - \mathbf{1}_A \otimes \log \xi_B)$  (see Corollary 7.3.8).

**Corollary 7.3.2.** *Given a state  $\rho_{AB} \in \mathcal{D}_+(\mathcal{H}_A \otimes \mathcal{H}_B)$  and  $\varepsilon > 0$ , a state  $\xi_{AB} \in B_\varepsilon^+(\rho_{AB})$  has maximum conditional entropy if and only if*

$$S(A|B)_\xi - S(A|B)_\rho + D(\rho_{AB} \| \xi_{AB}) - D(\rho_{AB} \| \xi_B) = \varepsilon(\lambda_+(L_\xi) - \lambda_-(L_\xi))$$

where  $L_\xi = \mathbf{1}_A \otimes \log \xi_B - \log \xi_{AB}$ .

We first prove the inequality (7.6) by considering the Jordan decomposition of  $\Delta := \xi - \rho$ . Next, we convert the constrained optimization problem  $\max_{\xi \in B_\varepsilon(\rho)} f(\xi)$  to an unconstrained optimization problem  $\min h$  for a non-Gâteaux differentiable function  $h$  defined on all of  $\mathcal{B}_{\text{sa}}(\mathcal{H})$  by adding appropriate indicator functions for the sets  $\mathcal{D}(\mathcal{H})$  and  $\{A \in \mathcal{B}_{\text{sa}}(\mathcal{H}) : \frac{1}{2}\|A - \rho\|_1 \leq \varepsilon\}$ .

The Moreau-Rockafellar Theorem (Theorem 7.1.3) allows us to compute  $\partial h(\xi)$  in terms of  $\partial f(\xi) = \{\nabla f(\xi)\}$  and the subgradients of the indicator functions. We then show  $0 \in \partial h(\xi)$  if and only if equality is achieved in (7.6).

**Proof of Theorem 7.3.1** We start with the following general result, which does not use convex optimization.

**Lemma 7.3.3.** *Let  $\varepsilon > 0$  and  $\Delta \in \mathcal{B}_{\text{sa}}$  with  $\text{tr}(\Delta) = 0$ ,  $\frac{1}{2}\|\Delta\|_1 \leq \varepsilon$ . Let  $L \in \mathcal{B}_{\text{sa}}$ . Then*

$$\text{tr}(L\Delta) \leq \varepsilon(\lambda_+(L) - \lambda_-(L)) \quad (7.7)$$

*with equality if and only if*

1. *Either  $\frac{1}{2}\|\Delta\|_1 = \varepsilon$  or else  $L = \lambda \mathbf{1}$  so that  $\lambda := \lambda_+(L) = \lambda_-(L)$ , and*
2. *we have*

$$\pi_{\pm} L \pi_{\pm} = \lambda_{\pm}(L) \pi_{\pm}$$

*where  $\pi_{\pm}$  is the projection onto the support of  $\Delta_{\pm}$ , and where  $\Delta = \Delta_+ - \Delta_-$  is the Jordan decomposition of  $\Delta$ .*

*Proof.* Note  $|\Delta| = \Delta_+ + \Delta_-$ , and since  $\Delta$  has trace zero,  $\text{tr}(\Delta_+) = \text{tr}(\Delta_-)$ . We expand

$$\text{tr}(L\Delta) = \text{tr}(L\Delta_+) - \text{tr}(L\Delta_-).$$

Now, we use that  $L$  is self-adjoint so that

$$\lambda_-(L) \mathbf{1} \leq L \leq \lambda_+(L) \mathbf{1}.$$

Since  $\Delta_{\pm} \geq 0$ , we have

$$\text{tr}(L\Delta_+) - \text{tr}(L\Delta_-) \leq \text{tr}(L\Delta_+) - \text{tr}(\lambda_-(L)\Delta_-) \quad (7.8)$$

$$\leq \text{tr}(\lambda_+(L)\Delta_+) - \text{tr}(\lambda_-(L)\Delta_-) \quad (7.9)$$

$$= (\lambda_+(L) - \lambda_-(L)) \text{tr}(\Delta_+),$$

where the last line follows from the fact that  $\text{tr}(\Delta_+) = \text{tr}(\Delta_-)$ . However,  $\text{tr}(\Delta_+) \leq \varepsilon$ , so we have

$$(\lambda_+(L) - \lambda_-(L)) \text{tr}(\Delta_+) \leq \varepsilon(\lambda_+(L) - \lambda_-(L)). \quad (7.10)$$

Thus, (7.7) follows. To obtain equality in (7.7), we require equality in (7.8), (7.9) and (7.10). Equality in (7.10) is equivalent to condition 1 in the statement of the lemma. We now show that equality in (7.8) and (7.9) is equivalent to condition 2.

Set  $\lambda_+ = \lambda_+(L)$ . Then since  $\pi_+$  is the projection onto the support of  $\Delta_+$ , we have  $\Delta_+ = \pi_+ \Delta_+ \pi_+$ . Using cyclicity of the trace, we obtain

$$\text{tr}(L\Delta_+) = \text{tr}(L\pi_+ \Delta_+ \pi_+) = \text{tr}(\pi_+ L \pi_+ \Delta_+).$$

Equality in (7.9) implies  $\text{tr}(L\Delta_+) = \lambda_+ \text{tr}(\Delta_+)$ , so

$$\text{tr}((\pi_+ L \pi_+) \Delta_+) = \text{tr}(\lambda_+ \Delta_+)$$

and thus

$$\text{tr}(\Delta_+(\lambda_+ \pi_+ - \pi_+ L \pi_+)) = 0. \quad (7.11)$$

Since  $\lambda_+$  is the largest eigenvalue of  $L$  which is self-adjoint, we have  $L \leq \lambda_+ \mathbf{1}$ . Since conjugating by any operator preserves the semidefinite order, we have  $\lambda_+ \pi_+ - \pi_+ L \pi_+ \geq 0$ . Then since  $\Delta_+$  restricted to  $\pi_+$  is positive definite, (7.11) implies  $\pi_+ L \pi_+ = \lambda_+ \pi_+$ . Similarly, equality in (7.8) implies  $\pi_- L \pi_- = \lambda_- (L) \pi_-$ .

Conversely, if  $\pi_\pm L \pi_\pm = \lambda_\pm \pi_\pm$ , we immediately obtain equality in (7.8) and (7.9).  $\square$

This lemma with the choices  $\Delta = \xi - \rho$  and  $L \equiv L_\xi := \nabla f(\xi)$  gives the inequality (7.6) of Theorem 7.3.1. It also gives the equivalence between the conditions 1 and 3 of the theorem.

We now turn to the theory of convex optimization to establish the equivalence between conditions 1 and 2. Recall  $f : \mathcal{D} \rightarrow \mathbb{R}$  is a continuous, concave function which is Gâteaux-differentiable on  $\mathcal{D}_+$ . Let us write  $\tilde{f} = -f$  which is convex, and note  $L_\xi = -\nabla \tilde{f}(\xi)$ . With this choice, it remains to be shown that  $\xi^* \in \text{argmin}_{\xi \in B_\varepsilon(\rho)} \tilde{f}(\xi)$  if and only if

$$\text{tr}(L_{\xi^*}(\xi^* - \rho)) \geq \varepsilon(\lambda_+(L_{\xi^*}) - \lambda_-(L_{\xi^*})). \quad (7.12)$$

The Tietze extension theorem (e.g. [Sim15, Theroem 2.2.5]) allows one to extend a bounded continuous function defined on a closed set of a normal topological space (such as a normed vector space) to the entire space, while preserving continuity and boundedness. We use this to extend  $\tilde{f}$  (which is bounded as it is a continuous function on the compact set  $\mathcal{D}$ ) to the whole space  $\mathcal{B}_{\text{sa}}$ , using that  $\mathcal{D}$  is a closed set in  $\mathcal{B}_{\text{sa}}$ . We consider the closed and convex sets  $\mathcal{D}$  and

$$T := \{A \in \mathcal{B}_{\text{sa}} : \|A - \rho_{AB}\|_1 \leq 2\varepsilon\}, \quad (7.13)$$

and note  $B_\varepsilon = \mathcal{D} \cap T$ . We define  $h : \mathcal{B}_{\text{sa}} \rightarrow \mathbb{R} \cup \{\infty\}$  as

$$h = \tilde{f} + \delta_{\mathcal{D}} + \delta_T$$

where for  $S \subseteq \mathcal{B}_{\text{sa}}$ , the indicator function  $\delta_S(A) = \begin{cases} 0 & A \in S \\ +\infty & \text{otherwise.} \end{cases}$

We have the important fact that, by construction,

$$\operatorname{argmin}_{\omega \in B_\varepsilon(\rho)} \tilde{f}(\omega) = \operatorname{argmin}_{A \in \mathcal{B}_{\text{sa}}} h(A). \quad (7.14)$$

By Theorem 7.1.2  $\hat{A} \in \mathcal{B}_{\text{sa}}$  is a global minimizer of  $h$  if and only if  $0 \in \partial h(\hat{A})$ . Note each of  $\tilde{f}$ ,  $\delta_{\mathcal{D}}$ ,  $\delta_T$  is lower semicontinuous, convex, and has non-empty domain. Moreover, both  $\tilde{f}$  and  $\delta_T$  are continuous at any faithful state  $\omega \in B_\varepsilon^+(\rho) \subseteq \operatorname{dom} \delta_{\mathcal{D}}$ . Hence, by Theorem 7.1.3,

$$\partial h = \partial \tilde{f} + \partial \delta_{\mathcal{D}} + \partial \delta_T := \{\ell_f + \ell_{\mathcal{D}} + \ell_T : \ell_f \in \partial \tilde{f}, \ell_{\mathcal{D}} \in \partial \delta_{\mathcal{D}}, \ell_T \in \partial \delta_T\}.$$

Hence, to obtain a complete characterization of

$$\operatorname{argmin} h = \{A \in \mathcal{B}_{\text{sa}} : 0 \in \partial h(A)\}$$

we need to evaluate the subgradients of  $\tilde{f}$ ,  $\delta_T$ , and  $\delta_{\mathcal{D}}$ .

Since  $\tilde{f}$  is Gâteaux-differentiable on  $\mathcal{D}_+$ , for any  $\omega \in \mathcal{D}_+$ , we have  $\partial \tilde{f}(\omega) = \{-L_\omega\}$  for  $L_\omega := -\nabla \tilde{f}(\omega)$ . The following two results (proven in Section 7.4) describe the other two relevant subgradients.

**Lemma 7.3.4.** *For  $A \in \mathcal{B}_{\text{sa}}$  with  $0 < \frac{1}{2}\|A - \rho_{AB}\|_1 \leq \varepsilon$ ,*

$$\partial \delta_T(A) = \{u \in \mathcal{B}_{\text{sa}} : 2\varepsilon\|u\|_\infty = \langle u, A - \rho_{AB} \rangle\}$$

where

$$\|u\|_\infty := \lambda_+(|u|) = \sup_{A \in \mathcal{B}_{\text{sa}}, \|A\|_1=1} |\langle u, A \rangle|.$$

**Lemma 7.3.5.** *Let  $\omega \in \mathcal{D}_+$ . Then  $\partial \delta_{\mathcal{D}}(\omega) = \{x\mathbf{1} : x \in \mathbb{R}\}$ .*

Putting together these results, we have, for  $\xi^* \in B_\varepsilon^+(\rho)$ ,

$$0 \in \partial h(\xi^*) \iff 0 = -L_{\xi^*} + G + x\mathbf{1}$$

for some  $G \in \partial \delta_T(\xi^*)$  and  $x \in \mathbb{R}$ , where  $G$  satisfies

$$2\varepsilon\|G\|_\infty = \operatorname{tr}(G(\xi^* - \rho)).$$

Then  $G = L_{\xi^*} - x\mathbf{1}$ , which implies that

$$\operatorname{tr}(L_{\xi^*}(\xi^* - \rho)) = 2\varepsilon\|L_{\xi^*} - x\mathbf{1}\|_\infty.$$

Set  $\alpha = \frac{\text{tr}(L_{\xi^*}(\xi^* - \rho))}{2\varepsilon}$ . Note  $\alpha$  and  $L_{\xi^*}$  depend on  $\xi^*$ . Then we have

$$0 \in \partial h(\xi^*) \iff \exists x \in \mathbb{R} : \alpha = \|L_{\xi^*} - x\mathbf{1}\|_{\infty}.$$

Since  $\|L_{\xi^*} - x\mathbf{1}\|_{\infty} = \max_{\lambda \in \text{spec } L_{\xi^*}} |\lambda - x|$ , we have that  $0 \in \partial h(\xi^*)$  if and only if

$$\exists x \in \mathbb{R} : \alpha = \max_{\lambda \in \text{spec } L_{\xi^*}} |\lambda - x|. \quad (7.15)$$

Let

$$q(x) = \max_{\lambda \in \text{spec } L_{\xi^*}} |\lambda - x| \quad (7.16)$$

We immediately see that  $q$  is continuous,  $q(0) = \lambda_+(|L_{\xi^*}|)$ , and  $\lim_{z \rightarrow \pm\infty} q(z) = +\infty$ . In fact, we can write a simple form for  $q(x)$  as the following lemma, which is proven in Section 7.4, shows. Set  $\lambda_+ \equiv \lambda_+(L_{\xi^*})$  and  $\lambda_- \equiv \lambda_-(L_{\xi^*})$  in the following.

**Lemma 7.3.6.** *The quantity  $q(x)$  defined by (7.16) can be expressed as follows:*

$$q(x) = \frac{\lambda_+ - \lambda_-}{2} + \left| \frac{\lambda_+ + \lambda_-}{2} - x \right|.$$

Thus, Lemma 7.3.6 implies that the range of the function  $q$  is  $[\frac{\lambda_+ - \lambda_-}{2}, \infty)$ . Hence, (7.15) holds if and only if

$$\frac{\lambda_+ - \lambda_-}{2} \leq \alpha.$$

Substituting  $\alpha = \frac{\text{tr}(L_{\xi^*}(\xi^* - \rho))}{2\varepsilon}$  in the above expression yields (7.12) and therefore concludes the proof of Theorem 7.3.1.  $\square$

Below, we collect some results (proven in Section 7.4) relating to the Gâteaux gradients of relevant entropic functions which are candidates for  $f$  in Theorem 7.3.1.

**Lemma 7.3.7.** *The von Neumann entropy  $\xi \mapsto S(\xi) := -\text{tr}(\xi \log \xi)$  is Gâteaux-differentiable at each  $\xi > 0$  and  $\nabla S(\xi) = -\log \xi - \frac{1}{\log_e(2)}\mathbf{1}$ .*

**Corollary 7.3.8.** *The conditional entropy  $\xi_{AB} \mapsto S(A|B)_{\xi} := S(\xi_{AB}) - S(\xi_B)$  is Gâteaux-differentiable at each  $\xi_{AB} > 0$  and  $\nabla S(A|B)(\xi_{AB}) = -(\log \xi_{AB} - \mathbf{1}_A \otimes \log \xi_B)$ .*

**Lemma 7.3.9.** *For  $\alpha \in (0, 1) \cup (1, \infty)$ , the map  $T_{\alpha}$  and the Rényi entropy  $S_{\alpha}$  are Gâteaux-differentiable at each  $\xi \in \mathcal{D}_+$ , with Gâteaux gradients*

$$\nabla T_{\alpha}(\xi) = \alpha \xi^{\alpha-1}, \quad \nabla S_{\alpha}(\xi) = \frac{\alpha}{1 - \alpha} \frac{1}{\ln 2} \frac{1}{\text{tr}[\xi^{\alpha}]} \xi^{\alpha-1}.$$

## 7.4 Proofs of the remaining lemmas and corollaries

**Proof of Corollary 7.3.2** We write the conditional entropy as the map

$$\begin{aligned} S(A|B)_{\cdot} &: \mathcal{D}(\mathcal{H}_{AB}) \rightarrow \mathbb{R} \\ \xi &\mapsto S(A|B)_{\xi}. \end{aligned} \quad (7.17)$$

By Corollary 7.3.8, for any  $\xi_{AB} > 0$  we have

$$\nabla S(A|B)(\xi) = -(\log \xi_{AB} - \mathbf{1}_A \otimes \log \xi_B) = -L_{\xi}.$$

Substituting this in the left-hand side of (7.6), we obtain

$$\begin{aligned} \text{tr}(L_{\xi}(\rho_{AB} - \xi_{AB})) &= \text{tr}[\rho_{AB} \log \xi_{AB}] - \text{tr}[\xi_{AB} \log \xi_{AB}] - \text{tr}[\rho_B \log \xi_B] + \text{tr}[\xi_B \log \xi_B] \\ &= S(A|B)_{\xi} + \text{tr}[\rho_{AB} \log \rho_{AB}] - \text{tr}[\rho_{AB}(\log \rho_{AB} - \log \xi_{AB})] \\ &\quad - \text{tr}[\rho_B \log \rho_B] + \text{tr}[\rho_B(\log \rho_B - \log \xi_B)] \\ &= S(A|B)_{\xi} - S(A|B)_{\rho} - D(\rho_{AB} \|\xi_{AB}) + D(\rho_B \|\xi_B). \end{aligned}$$

The statement of Corollary 7.3.2 then follows from Theorem 7.3.1.  $\square$

**Proof of Lemma 7.3.4** Let  $\mathcal{A} := \{u \in \mathcal{B}_{\text{sa}} : 2\varepsilon\|u\|_{\infty} = \langle u, A - \rho_{AB} \rangle\}$  and set

$$C := \{B \in \mathcal{B}_{\text{sa}} : \|B - \rho_{AB}\|_1 \leq \varepsilon\} \subseteq \mathcal{B}_{\text{sa}}.$$

Let  $u \in \mathcal{A}$ . Then for  $y \in C$ ,

$$\begin{aligned} \langle u, y - A \rangle &= \langle u, y - \rho_{AB} \rangle + \langle u, \rho_{AB} - A \rangle = \langle u, y - \rho_{AB} \rangle - 2\varepsilon\|u\|_{\infty} \\ &\leq \|u\|_{\infty}\|y - \rho_{AB}\|_1 - 2\varepsilon\|u\|_{\infty} \\ &= \|u\|_{\infty}(\|y - \rho_{AB}\|_1 - 2\varepsilon) \leq 0 \end{aligned}$$

and thus  $u \in \partial\delta_T(A)$ . On the other hand, take  $u \in \partial\delta_T(A)$ . Then

$$\|u\|_{\infty} = \sup_{\|x\|_1=1} |\langle u, x \rangle|$$

is achieved at some  $x^* \in \mathcal{B}_{\text{sa}}$  since the set  $\{x \in \mathcal{B}_{\text{sa}} : \|x\|_1 = 1\}$  is compact, using that  $\mathcal{B}_{\text{sa}}$  is finite-dimensional. Then  $y = \rho_{AB} + 2\varepsilon x \in C$ . Hence,

$$0 \geq \langle u, y - A \rangle = \langle u, \rho_{AB} - A \rangle + 2\varepsilon \langle u, x \rangle = \langle u, \rho_{AB} - A \rangle + 2\varepsilon\|u\|_{\infty}$$

and thus  $\langle u, A - \rho_{AB} \rangle \geq 2\varepsilon \|u\|_\infty$ . Then by the bound

$$\langle u, A - \rho_{AB} \rangle \leq \|u\|_\infty \|A - \rho_{AB}\|_1 \leq 2\varepsilon \|u\|_\infty$$

we have  $u \in \mathcal{A}$ . □

**Proof of Lemma 7.3.5** By definition,

$$\partial\delta_{\mathcal{D}}(\omega) = \{u \in \mathcal{B}_{\text{sa}} : \langle u, y - \omega \rangle \leq 0 \text{ for all } y \in \mathcal{D}\}.$$

Since  $u$  is self-adjoint, we can write its eigen-decomposition as  $u = \sum_{k=1}^d \alpha_k |k\rangle\langle k|$ . If  $\alpha_k = \alpha_j$  for each  $j, k$ , then  $u \propto \mathbf{1}$ . Conversely,  $\alpha_k \mathbf{1} \in \partial\delta_{\mathcal{D}}(\omega)$  since  $\text{tr}[\alpha_k \mathbf{1}(y - \omega)] = \alpha_k(\text{tr}(y) - \text{tr}(\omega)) = 0$ .

Otherwise, assume  $\alpha_k > \alpha_j$  for some  $k, j \in \{1, \dots, d\}$ . Let

$$y := \sum_{i, i \neq k, i \neq j} \langle i|\omega|i\rangle |i\rangle\langle i| + (\langle k|\omega|k\rangle + \langle j|\omega|j\rangle) |k\rangle\langle k| \in \mathcal{B}_{\text{sa}}.$$

Then  $\text{tr}(y) = \sum_i \langle i|\omega|i\rangle = \text{tr}(\omega) = 1$ , and  $y \geq 0$  since all of its eigenvalues are non-negative. Note  $|j\rangle$  is an eigenvector of  $y$  with eigenvalue zero. Next,

$$\begin{aligned} \text{tr}(u(y - \omega)) &= \sum_i \langle i|(y - \omega)|i\rangle \\ &= \sum_{i, i \neq k, i \neq j} (\langle i|\omega|i\rangle - \langle i|\omega|i\rangle) \\ &\quad + \alpha_k(\langle k|\omega|k\rangle + \langle j|\omega|j\rangle - \langle k|\omega|k\rangle) + \alpha_j(0 - \langle j|\omega|j\rangle) \\ &= (\alpha_k - \alpha_j) \langle j|\omega|j\rangle > 0. \end{aligned}$$

Thus,  $u \notin \partial\delta_{\mathcal{D}}(\omega)$ . □

**Proof of Lemma 7.3.6** First, we establish

$$q(x) = \max(|\lambda_+ - x|, |\lambda_- - x|). \quad (7.18)$$

Given  $\lambda \in \text{spec } L$  and  $x \in \mathbb{R}$ , if  $x < \lambda$ , then

$$|x - \lambda| = \lambda - x \leq \lambda_+ - x = |\lambda_+ - x|$$

and otherwise

$$|x - \lambda| = x - \lambda \leq x - \lambda_- = |\lambda_- - x|,$$

yielding (7.18). Next, set  $r := \frac{\lambda_+ - \lambda_-}{2}$  and  $m := \frac{\lambda_+ + \lambda_-}{2}$ . Then  $\lambda_{\pm} = m \pm r$ . Therefore, for  $x \in \mathbb{R}$ ,

$$|\lambda_{\pm} - x| = |m \pm r - x| \leq r + |m - x|,$$

yielding  $q(x) = r + |m - x|$  as claimed.  $\square$

**Proof of Lemma 7.3.7** Let us introduce the Cauchy integral representation of an analytic function. If  $g$  is analytic on a domain  $G \subseteq \mathbb{C}$  and  $A$  is a matrix with  $\text{spec } A \subseteq G$ , then we can write

$$g(A) = \frac{1}{2i\pi} \int_{\Gamma} g(z)(z\mathbf{1} - A)^{-1} dz,$$

and

$$g'(A) = \frac{1}{2i\pi} \int_{\Gamma} g(z)(z\mathbf{1} - A)^{-2} dz, \quad (7.19)$$

where  $\Gamma \subseteq G$  is a simple closed curve with  $\text{spec } A \subseteq \Gamma$ , and  $g' : G \rightarrow \mathbb{C}$  is the derivative of  $g$ . [Sti87] shows that with these definitions, the Fréchet derivative of  $g$  at  $A$  exists, and when applied to a matrix  $X$  yields

$$D(g)(A)X = \frac{1}{2i\pi} \int_{\Gamma} g(z)(z\mathbf{1} - A)^{-1} X (z\mathbf{1} - A)^{-1} dz.$$

Therefore, by cyclicity of the trace,

$$\text{tr}(D(g)(A)X) = \text{tr} \left[ \frac{1}{2i\pi} \int_{\Gamma} g(z)(z\mathbf{1} - A)^{-2} dz X \right] = \text{tr}(g'(A)X) \quad (7.20)$$

using (7.19) for the second equality.

We may write the von Neumann entropy as the map

$$S : \mathcal{D} \rightarrow \mathbb{R}, \quad \xi \mapsto S(\xi) = -\text{tr}[\xi \log \xi].$$

Then we may write  $S = \text{tr} \circ \eta$  for  $\eta(x) = -x \log x$  which is analytic on  $\{\zeta \in \mathbb{C} : \text{Re } \zeta > 0\}$ , with derivative  $\eta'(x) = -\log x - \frac{1}{\log_e(2)}$ . Then for  $\xi > 0$ ,

$$D(S)(\xi) = D(\text{tr} \circ \eta)(\xi) = D(\text{tr})(\eta(\xi)) \circ D(\eta)(\xi).$$



by the chain rule for Fréchet derivatives. By the linearity of the trace,  $D(\text{tr})B(X) = \text{tr}(X)$  for any  $X, B \in \mathcal{B}_{\text{sa}}$ . So for  $X \in \mathcal{B}_{\text{sa}}$ ,

$$D(S)(\xi)X = \text{tr}[D(\eta)(\xi)X].$$

By (7.20) for  $g = \eta$ , we have

$$D(S)(\xi)X = \text{tr}[\eta'(\xi)X] = \text{tr}\left[\left(-\log \xi - \frac{1}{\log_e(2)}\mathbf{1}\right)X\right]. \quad \square$$

**Proof of Corollary 7.3.8** We decompose the conditional entropy map (Equation (7.17)) by writing

$$S(A|B) = S_{AB}(\cdot) - S_B \circ \text{tr}_A(\cdot) : \mathcal{D}(H_{AB}) \rightarrow \mathbb{R}$$

where we have indicated explicitly the domain of each function in the subscript. That is,  $S_{AB} : \mathcal{D}(\mathcal{H}_{AB}) \rightarrow \mathbb{R}$  is the von Neumann entropy on system  $AB$ , and  $S_B : \mathcal{D}(\mathcal{H}_B) \rightarrow \mathbb{R}$  is the von Neumann entropy on  $B$ .

Since  $\text{tr}_A : \mathcal{B}_{\text{sa}}(\mathcal{H}_{AB}) \rightarrow \mathcal{B}_{\text{sa}}(\mathcal{H}_B)$  is a bounded linear map and  $S_B$  concave and continuous, the chain rule for the composition with a linear map (see Prop. Prop. 3.28 of [Pey15]) gives

$$\nabla(S_B \circ \text{tr}_A)(\omega_{AB}) = \text{tr}_A^* \circ \nabla S_B(\omega_B) = \mathbf{1}_A \otimes \nabla S_B(\omega_B)$$

for any  $\omega_{AB} \in \mathcal{D}(\mathcal{H}_{AB})$ , where  $\text{tr}_A^*$  is the dual to the map  $\text{tr}_A$ . As  $\nabla S_B(\omega_B) = -(\log \omega_B + \frac{1}{\log_e(2)}\mathbf{1}_B)$ , we have

$$\begin{aligned} \nabla S(A|B)(\omega_{AB}) &= \nabla S_{AB}(\omega_{AB}) - \nabla(S_B \circ \text{tr}_A)(\omega_{AB}) \\ &= -(\log \omega_{AB} + \frac{1}{\log_e(2)}\mathbf{1}_{AB}) + \mathbf{1}_A \otimes (\log \omega_B + \frac{1}{\log_e(2)}\mathbf{1}_B) \\ &= -(\log \omega_{AB} - \mathbf{1}_A \otimes \log \omega_B). \quad \square \end{aligned}$$

**Proof of Lemma 7.3.9** The  $\alpha$ -Rényi entropy  $\alpha \in (0, 1) \cup (1, \infty)$  can be described by the map

$$S_\alpha : \mathcal{D} \mapsto \mathbb{R}, \quad \xi \mapsto S_\alpha(\xi) := \frac{1}{1-\alpha} \log \text{tr}(\xi^\alpha).$$

Let us use the notation  $\text{pow}_\alpha : \mathcal{B}_{\text{sa}} \rightarrow \mathcal{B}_{\text{sa}}$ ,  $A \mapsto \text{pow}_\alpha(A) := A^\alpha$ . Then

$$T_\alpha = \text{tr} \circ \text{pow}_\alpha$$

and  $(1-\alpha)S_\alpha = \log \circ T_\alpha$ . For  $\xi \in \mathcal{D}_+$ , the function  $\text{pow}_\alpha$  is analytic on  $\{\zeta \in \mathbb{C} : \text{Re } \zeta > 0\}$ . As in the proof of Lemma 7.3.7 then, using the chain rule and linearity of the trace we find

$$D(T_\alpha)(\xi) = \text{tr}[D(\text{pow}_\alpha)(\xi)X].$$

Invoking (7.20) for  $g(z) = z^\alpha$ , with  $\Gamma \subseteq \{\zeta \in \mathbb{C} : \text{Re } \zeta > 0\}$  a simple closed curve enclosing the spectrum of  $\xi$ , we have

$$D(T_\alpha)(\xi)X = \alpha \text{tr}(\xi^{\alpha-1}X). \quad (7.21)$$

Moreover,

$$\begin{aligned} D(S_\alpha)(\xi)X &= \frac{1}{1-\alpha} D(\log \circ T_\alpha)(\xi) \\ &= \frac{1}{1-\alpha} D(\log)(\text{tr}(\xi^\alpha)) \circ D(T_\alpha)(\xi) \\ &= \frac{1}{1-\alpha} \frac{1}{\ln 2} \frac{\alpha}{\text{tr}[\xi^\alpha]} \text{tr}(\xi^{\alpha-1}X). \quad \square \end{aligned}$$

# Chapter 8

## Open questions from Part I

1. Section 3.7 shows that balls in  $p$ -norm with  $1 < p < \infty$  do not admit majorization-minimizers, while Chapter 3 is dedicated to proving that balls in 1-norm do admit majorization-minimizers (and majorization-maximizers). Do balls in Hellinger distance, Bures distance, the (square-root) Shannon-Jensen divergence, Wasserstein distances, or other metrics on  $\mathcal{P}$  admit majorization-minimizers?
2. In Chapter 4, the notion of majorization flow is defined in terms of the majorization-minimizer. Can a similar notion be defined in terms of the majorization-maximizer? Would it be useful?
3. In Chapter 5, the path of majorization flow (defined in Chapter 4) is used for understanding the continuity properties of various Schur concave functions. Does it have other applications?
4. Theorem 5.1.1 shows that Concave-Type  $(h, \phi)$ -entropies  $H$ , which are symmetric and concave, are such that their derivative along the path of majorization flow,  $\Gamma_H$ , is strictly Schur convex. Proposition 5.1.6 shows that there exist symmetric and concave functions  $f$  such that  $\Gamma_f$  is not Schur convex. Is there a simple characterization of functions  $f : \mathcal{P} \rightarrow \mathbb{R}$  such that  $\Gamma_f$  is Schur convex?
5. Proposition 5.2.2 provides upper and lower bounds on the optimal Lipschitz constant for the  $\alpha$ -Rényi entropy on  $\mathcal{P}_d$  for  $\alpha > 1$ , and an exact expression for  $\alpha \in \{2, \infty\}$ . The lower bound scales with  $d$  as  $(d - 2)^{1-1/\alpha}$  while the upper bound scales as  $d$ . What is the correct scaling for  $\alpha \in (1, 2) \cup (2, \infty)$ ?

6. Proposition 5.7.1 shows that the majorization-minimizer and -maximizer can be used to provide an approximate source or target state for LOCC conversions of bipartite pure states. Are these states optimal with respect to some criterion?
7. Proposition 6.3.3 provides upper and lower bounds on the optimal Lipschitz constant for the guesswork with quantum side information which match up to constant factors. What is the exact optimal Lipschitz constant?
8. Section 6.3.6 provides entropic bounds on the guesswork with quantum side information. Is there a single-letter expression for the asymptotic quantity  $\lim_{n \rightarrow \infty} \frac{1}{n} G(X^n | B^n)_{\rho^{\otimes n}}$ ?
9. Corollary 7.3.2 presents necessary and sufficient conditions for a state to maximize the conditional entropy over a ball in trace distance around a given state. Can these conditions be “inverted” to obtain an expression for the maximizer, or the maximum value of the conditional entropy? See also [AS19; Wil20] for recent progress on uniform continuity bounds for the conditional entropy.

## Part II

# Eventually Entanglement-Breaking Channels



# Chapter 9

## Introduction

In this part of the thesis, we study the behavior of the composition of quantum channels  $\Phi_n \circ \Phi_{n-1} \circ \dots \circ \Phi_1$  in two settings.

In the first setting, discussed in Chapter 11, the channels are all identical, and we investigate for which channels  $\Phi$  there exists  $n \in \mathbb{N}$  such that  $\Phi^n := \underbrace{\Phi \circ \dots \circ \Phi}_{n \text{ times}}$  is *entanglement-breaking*, meaning that when applied to a state, any initially present entanglement with a reference system is destroyed. Such channels are called *eventually entanglement-breaking* (EEB), and provide the title of this part of the thesis. Under the assumption that  $\Phi$  admits a full-rank invariant state (i.e. that  $\Phi$  is *faithful*), we characterize which channels  $\Phi$  are eventually entanglement-breaking in Theorem 11.3.8.

Let us briefly discuss the relevance of this result to quantum information theory. First, many quantum information-theoretic protocols require entangled qubits (ebits) shared between separated parties Alice and Bob as a resource. However, if Alice or Bob's laboratory is subject to local noise, initially entangled states may soon become unentangled, and Alice and Bob may not be able to carry out the protocol. As quantum channels can be seen as a model of noise, characterizing which channels are eventually-entanglement breaking provides a means to understand which types of noise lead to complete loss of entanglement in finite time.

The noise model presented by an  $n$ -fold composition  $\Phi^n$  is not generic, however, but instead possesses a discrete-time Markovianity, represented by the semi-group property  $\Phi^{n+m} = \Phi^n \circ \Phi^m$ . This can be given a physical interpretation as follows. The Stinespring dilation theorem states that any quantum channel  $\Phi$ , the model for general time evolution of a quantum system  $\mathcal{S}$ , can be written as

$$\Phi(\rho_{\mathcal{S}}) = \text{tr}_{\mathcal{E}}[U_{\mathcal{SE}}(\rho_{\mathcal{S}} \otimes \psi_{\mathcal{E}})U_{\mathcal{SE}}^*]$$

for some environment system  $\mathcal{E}$ , joint unitary  $U_{\mathcal{SE}}$ , and pure state  $\psi_{\mathcal{E}}$ . The time evolution of  $\rho_{\mathcal{S}}$ , therefore, can be represented by a noiseless interaction with an environment system  $\mathcal{E}$ , followed by the loss of  $\mathcal{E}$  itself, inducing noise on the system. The composition  $\Phi^n$ , therefore, models an evolution of  $\rho_{\mathcal{S}}$  which consists of  $n$  consecutive identical lossless interactions with separate environment systems, followed by the loss of each. Thus, between subsequent interactions the environment does not “remember” the details of the previous interaction; instead, a fresh environment is procured.

This Markovianity aids in the mathematical analysis of the evolution, and provides a good representation of some physical systems, such as the Haroche experiments [GKGB+07], but is not a faithful model of all types of time evolution. This model is, however, more general than the oft-considered model provided by a continuous time quantum Markov semigroup, in which time evolution from 0 to  $t \in \mathbb{R}_{>0}$  is given by  $\Phi_t = \exp(t\mathcal{L})$  for a Lindblad generator  $\mathcal{L}$ . The reduction follows from the fact that  $\Phi_1$  is a quantum channel such that  $\Phi_{n+m} = \Phi_1^{n+m}$ . Note that while the continuous-time model can represent non-integer times as well, for the purposes of determining whether or not a channel is eventually entanglement-breaking, one can consider the decomposition  $\Phi_t = \Phi_{t-\lfloor t \rfloor} \circ \Phi_{\lfloor t \rfloor}$  to see that as soon as  $\Phi_{\lfloor t \rfloor}$  is entanglement-breaking,  $\Phi_t$  is too. On the other hand, not all quantum channels  $\Phi$  can be obtained in such a manner from a continuous-time quantum Markov semigroup; this is the topic of the divisibility of quantum channels, discussed in e.g. [WC08].

The second information-theoretic application of this result relates to quantum key repeaters introduced in [BCHW15]. Consider the task of distilling private states between Alice and Bob, who each individually share a state with an intermediary, Charlie: Alice shares  $n$  copies of state  $\rho$  with Charlie, and Bob shares  $n$  copies of a state  $\rho'$  with Charlie. Charlie acts as a “repeater”, facilitating communication between Alice and Bob. The three parties use a tripartite LOCC protocol to approximately recover  $nR$  private states shared between Alice and Bob; here,  $R$  is the *rate* of the task.

This task is related to the so-called PPT<sup>2</sup> conjecture [BKHW12], which conjectures that the composition of two PPT quantum channels (i.e. those whose Choi-Jamiolkowski state has a positive partial transpose) is entanglement-breaking. Translated to the language of bipartite states via the Choi-Jamiolkowski isomorphism (discussed in Section 10.2.2), this conjecture states that the quantum state obtained by entanglement swapping of PPT states (those with positive partial transpose) is separable. If true, the PPT<sup>2</sup> conjecture shows if  $\rho$  and  $\rho'$  are PPT states in the quantum key repeater task described above, the asymptotic rate is zero [BCHW15; CF17], in contrast to the task of



distilling private states from PPT states (without an intermediary), which can have a strictly positive rate [HHHO05].

Theorem 11.3.10, which can be seen as a corollary of the characterization of eventually entanglement-breaking channels discussed above, makes a step towards establishing the PPT<sup>2</sup> conjecture by showing that any faithful PPT completely positive map is eventually entanglement-breaking. This extends the recent results of [RJP18], which establish this statement for unital quantum channels (i.e. quantum channels such that the completely mixed state is invariant under the action of the channel).

The second setting, discussed in Chapter 12, is a slight generalization of the first, in which the maps  $\Phi_1, \Phi_2, \dots, \Phi_n$  are not necessarily identical, but have a particular extended form

$$\Phi_k(\rho_S) = \text{tr}_{\mathcal{E}_k} [U_{S\mathcal{E}_k}(\rho_S \otimes \xi_{\mathcal{E}_k})U_{S\mathcal{E}_k}^*]$$

in which the state of the environment,  $\xi_{\mathcal{E}_k}$ , is a thermal state instead of a pure state. This setup yields a *repeated interaction system* (or RIS), which is a system where the environment consists of a sequence of “probes”  $\mathcal{E}_k$ ,  $k = 1, \dots, T$ , initially in a thermal state at inverse temperature  $\beta_k$ , and  $\mathcal{S}$  interacts with  $\mathcal{E}_k$  (and only  $\mathcal{E}_k$ ) during the time interval  $[k\tau, (k+1)\tau)$  for some interaction time  $\tau > 0$ . Here,  $T \in \mathbb{N}$  is the total number of probes.

Chapter 12 discusses the results of [HJPR18], which studies a refinement of Landauer’s Principle in terms of a two-time measurement protocol (better known as “full counting statistics”) for repeated interaction systems, in an adiabatic regime. We briefly describe each of these elements below.

Landauer’s Principle is a universal principle for which the most common formulation is as a lower bound on the energetic cost of erasing a bit of information in a fixed system  $\mathcal{S}$  by interaction with an environment  $\mathcal{E}$  which is initially at thermal equilibrium. It was first stated by Landauer in [Lan61]. A recent, mathematically sound derivation (in [RW14], later extended to the case of infinitely extended systems in [JP14]) is based on the entropy balance equation, given by  $\Delta S_S + \sigma = \beta \Delta Q_{\mathcal{E}}$  where  $\Delta S_S$  is the average decrease in entropy of  $\mathcal{S}$  during the process,  $\Delta Q_{\mathcal{E}}$  the average increase in energy of  $\mathcal{E}$ , and  $\beta$  is the inverse temperature of the environment<sup>1</sup>. The term  $\sigma$  is called the *entropy production* of the process. As it can be written as a relative entropy, the entropy production is non-negative which yields the inequality  $\Delta S_S \leq \beta \Delta Q_{\mathcal{E}}$ . One of the questions of interest regarding Landauer’s Principle concerns the saturation of that identity, i.e. the vanishing of  $\sigma$ . It is a general physical principle that when the system–environment coupling is described by a time-dependent Hamiltonian, the entropy production  $\sigma$  vanishes in

<sup>1</sup>we will always set the Boltzmann constant to 1, so that  $\beta = 1/\Theta$ ,  $\Theta$  the temperature.

the adiabatic limit, that is, when the coupling between  $\mathcal{S}$  and  $\mathcal{E}$  is a slowly varying time-dependent function. More precisely, if the typical time scale of the coupling is  $T$ , one considers the regime  $T \rightarrow \infty$ .

For a repeated interaction system, the entropy balance equation becomes  $\sum_{k=1}^T \Delta S_{\mathcal{S}} + \sum_{k=1}^T \sigma_k = \sum_{k=1}^T \beta_k \Delta Q_{\mathcal{E},k}$ , where each term with index  $k$  corresponds to the interaction between  $\mathcal{S}$  and  $\mathcal{E}_k$ . We describe the repeated interaction system as an “adiabatic RIS” when the various parameters of the probes are sampled from sufficiently smooth functions on  $[0, 1]$  as the values at times  $k/T$ ,  $k = 1, \dots, T$ . This is the setup that was studied in [HJPR17]; there it was shown that the total entropy production  $\lim_{T \rightarrow \infty} \sum_{k=1}^T \sigma_k$  is finite only under the condition  $X(s) = 0$  for all  $s \in [0, 1]$ , where  $X(s)$  is a quantity depending on the probe parameters at time  $s \in [0, 1]$  discussed below (see (12.17)). The proof of this result relied mostly on a new discrete, non-unitary adiabatic theorem to estimate the behavior of  $\Phi_T \circ \dots \circ \Phi_2 \circ \Phi_1$ , where  $\Phi_1, \Phi_2, \dots, \Phi_T$  are slowly varying quantum channels that represent the reduced dynamics acting on  $\mathcal{S}$ .

A refinement of the above formulation of Landauer’s Principle is however possible using the so-called full counting statistics. Full counting statistics were first introduced in the study of charge transport, and have met with success in the study of fluctuation relations and work in quantum mechanics (see Kurchan [Kur00] and Tasaki [Tas00]). An example of their use in improving Landauer’s Principle was given in [BFJP17; GCGP+17]. In the present situation, the formulation of Landauer’s Principle in terms of full counting statistics can be stated by defining random variables  $\Delta s_{\mathcal{S}}$  and  $\Delta q_{\mathcal{E}_k}$  which are outcomes of simple physical experiments, which we now describe. In such an experiment, one initially measures the quantity  $-\log \rho_{\mathcal{S}}$  ( $\rho_{\mathcal{S}}$  is the state of the system) and the energies  $h_{\mathcal{E}_k}$  for each  $k$  ( $h_{\mathcal{E}_k}$  is the free Hamiltonian of  $\mathcal{E}_k$ ), then lets the system interact with the chain of probes, then measures again the same quantities. Note that it makes no difference if one measures each probe immediately after it interacts with the system, or measures all the probes at the end of the protocol. With the right sign conventions, the changes in these quantities are random variables which we denote  $\Delta s_{\mathcal{S}}$  and  $\Delta q_{\mathcal{E}_k}$ . Our refinement discusses the connections between the probability distributions of  $\Delta s_{\mathcal{S}}$  and  $\sum_k \beta_k \Delta q_{\mathcal{E}_k}$ . One can show that the expectations of these distributions are  $\Delta S_{\mathcal{S}}$  and  $\sum_k \beta_k \Delta Q_{\mathcal{E}_k}$  respectively; there is, therefore, more information in these distributions than in the previously considered scalar quantities.

We consider an adiabatic repeated interaction system and study the limiting distributions of the above random variables as  $T \rightarrow \infty$ . Again, we show that in the case  $X(s) \equiv 0$  we have the expected refinement of Landauer’s Principle, which is essentially that when  $T \rightarrow \infty$ , one has  $\Delta s_{\mathcal{S}} = \sum_k \beta_k \Delta q_{\mathcal{E}_k}$  almost-surely. In the case  $X(s) \not\equiv 0$ , we

show that  $\sum_k \beta_k \Delta q_{\mathcal{E}_k}$  satisfies a law of large numbers, a central limit theorem, and a large deviation principle, all of these for the time scale  $T$ , and with explicit parameters. In particular,  $\sum_k \beta_k \Delta q_{\mathcal{E}_k}$  is of order  $T$ , whereas  $\Delta s_S$  is a bounded quantity. All results in the case  $X(s) \not\equiv 0$  can actually be extended to the case where the probe observables measured at each step  $k$  are not simply  $\beta_k h_{\mathcal{E}_k}$  but a more general observable, or when the system observables are not  $-\log \rho_S$ .

We show in addition that the random variable  $\varsigma_T = \sum_k \beta_k \Delta q_{\mathcal{E}_k} - \Delta s_S$  can be expressed as a relative information random variable between the probability measure describing the experiment outcomes, and the probability measure corresponding to a backwards experiment. Since we obtain a full large deviation principle for this random variable as  $T \rightarrow \infty$ , this connects these results with the appearance of the arrow of time (see [ABL64; BJPP18]). We discuss in particular the appearance of symmetries in the moment generating functions, and their implications in terms of Gallavotti–Cohen type symmetries.

This approach gives an improvement over [HJPR17] in various aspects. First of all, Theorem 12.8.1 (in the case  $X(s) \equiv 0$ ) and Theorem 12.8.3 together with Corollary 12.8.4 and Theorem 12.8.5 (in the general case) characterize the limiting distributions of relevant random variables, whereas in [HJPR17] only information about the behaviour of their expectations was derived. We recover those former results (and more) about these expectations, as Theorem 12.8.1 implies in particular the convergence of  $\lim_{T \rightarrow \infty} \sum_{k=1}^T \sigma_k$  to an explicit quantity when  $X(s) \equiv 0$ , and Theorem 12.8.5 gives the divergence of the same quantity under generic assumptions when  $X(s)$  does not vanish identically. In addition, Corollary 12.7.2 gives an expression for the adiabatic evolution of any initial state.



# Chapter 10

## Linear Algebra and entanglement theory

In this chapter, we review fundamental areas of linear algebra and entanglement-theory used in the following chapters. Much of this chapter is written with an eye towards precision, at the cost of notational complexity, which is relaxed in the later chapters.

Let  $\mathcal{X}$  and  $\mathcal{Y}$  be finite-dimensional vector spaces. We write

- $\mathcal{B}(\mathcal{X} \rightarrow \mathcal{Y})$  for the set of linear maps from  $\mathcal{X}$  to  $\mathcal{Y}$
- $\mathcal{B}(\mathcal{X}) \equiv \mathcal{B}(\mathcal{X} \rightarrow \mathcal{X})$  for the set of linear operators on  $\mathcal{X}$
- $\mathcal{B}_{\text{sa}}(\mathcal{X})$  for the set of self-adjoint operators on  $\mathcal{X}$
- $\mathcal{B}_+(\mathcal{X}) \subseteq \mathcal{B}_{\text{sa}}(\mathcal{X})$  for the set of positive semidefinite operators on  $\mathcal{X}$ ,
- $\mathcal{B}_{++}(\mathcal{X}) \subseteq \mathcal{B}_+(\mathcal{X})$  for the set of full-rank positive semidefinite operators on  $\mathcal{X}$
- $\mathcal{D}(\mathcal{X})$  for the set of unit trace positive semidefinite operators, i.e. *density matrices* or *states*, on  $\mathcal{X}$ , and
- $\mathcal{D}_+(\mathcal{X}) \subseteq \mathcal{D}(\mathcal{X})$  for the set of full-rank density matrices. We say also call a state  $\sigma \in \mathcal{D}_+(\mathcal{X})$  a *faithful state*.

In the following,  $\mathcal{H}$ ,  $A$ ,  $A'$ ,  $B$ , and  $B'$  represent finite-dimensional Hilbert spaces, with dimensions  $d_{\mathcal{H}}$ ,  $d_A$ ,  $d_{A'}$ ,  $d_B$ , and  $d_{B'}$  respectively, and inner products  $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ ,  $\langle \cdot, \cdot \rangle_A$ ,  $\langle \cdot, \cdot \rangle_{A'}$ ,  $\langle \cdot, \cdot \rangle_B$ , and  $\langle \cdot, \cdot \rangle_{B'}$ . We often use  $\mathcal{H}$  in the case when there is only one Hilbert space under consideration.

We may equip  $\mathcal{B}(A \rightarrow B)$  with the Hilbert-Schmidt inner product  $\langle \cdot, \cdot \rangle_{\text{HS}}$  by

$$\langle X, Y \rangle_{\text{HS}} = \text{tr}[X^*Y] \quad (10.1)$$

where  $X^* \in \mathcal{B}(B \rightarrow A)$  is the adjoint of  $X$ , defined by

$$\langle \psi_B, X\phi_A \rangle_B = \langle X^*\psi_B, \phi_A \rangle_A \quad \forall \phi_A \in A, \psi_B \in B. \quad (10.2)$$

Then  $X^*Y \in \mathcal{B}(A \rightarrow A)$ , and  $\text{tr}[X^*Y]$  is defined as

$$\text{tr}[X^*Y] := \sum_{i=1}^{d_A} \langle \psi_i, X^*Y\psi_i \rangle_A$$

for any orthonormal basis  $\{|\psi_i\rangle\}_{i=1}^{d_A}$  of  $A$ .

In fact,  $\mathcal{X} := (\mathcal{B}(A \rightarrow B), \langle \cdot, \cdot \rangle_{\text{HS}})$  is itself a Hilbert space. Therefore, given  $\mathcal{X} := (\mathcal{B}(A \rightarrow B), \langle \cdot, \cdot \rangle_{\text{HS}})$  and  $\mathcal{Y} := (\mathcal{B}(C \rightarrow D), \langle \cdot, \cdot \rangle_{\text{HS}})$ , the space  $\mathcal{B}(\mathcal{X} \rightarrow \mathcal{Y})$  can be equipped with its own Hilbert-Schmidt inner product, and so forth. Therefore, for  $X, Y \in \mathcal{B}(A \rightarrow B)$ , we sometimes write  $\langle X, Y \rangle_{\mathcal{B}(A \rightarrow B), \text{HS}} \equiv \langle X, Y \rangle_{\text{HS}}$  for clarity.

## 10.1 Vectors, operators, and superoperators

A superoperator is a linear map  $\Phi \in \mathcal{B}(\mathcal{B}(A) \rightarrow \mathcal{B}(B))$ , where  $A$  and  $B$  are finite-dimensional Hilbert spaces. The space  $\mathcal{B}(\mathcal{B}(A) \rightarrow \mathcal{B}(B))$  has a lot of structure which we will explore in this section. It is helpful to start with the least structure and then see what each additional piece of structure adds to the theory. As a high-level overview, we will see that:

1. Vector spaces  $\mathcal{X}$  have scalar multiplication and vector addition, and admit direct sum decompositions. They can also be normed.
2. The set of linear operators on  $\mathcal{X}$ ,  $\mathcal{B}(\mathcal{X})$ , admits another operation, composition (i.e. vector multiplication), turning  $\mathcal{B}(\mathcal{X})$  into an algebra. Additionally, an operator in  $\mathcal{B}(\mathcal{X})$  can induce direct sum decompositions on  $\mathcal{X}$ , such as the spectral decomposition. As a vector space,  $\mathcal{B}(\mathcal{X})$  can be normed, and norms on  $\mathcal{X}$  induce operator norms on  $\mathcal{B}(\mathcal{X})$ .
3. Setting  $\mathcal{X} = A$ , a Hilbert space, adds an inner product to  $A$  and hence an adjoint to  $\mathcal{B}(\mathcal{X})$ , as well as a norm on  $A$  (the norm induced by the inner product). This gives rise to the adjoint and hence the notion of self-adjointness, as well as positive

semidefiniteness. Together,  $\mathcal{B}(A)$  with the operator norm  $\|\cdot\|_\infty$  induced by the Hilbert space norm on  $A$  and the adjoint  $*$  induced by the inner product on  $A$  yield a  $C^*$ -algebra (i.e. an algebra in which the norm, adjoint, and the composition operation satisfy key relations, such as the submultiplicativity of the norm, and the  $C^*$ -identity  $\|X^*X\| = \|X\|^2$  for  $X \in A$ ).

4. Raising one more level to the level of superoperators, i.e. the set  $\mathcal{B}(\mathcal{B}(A))$ , allows the consideration of which maps commute with the adjoint and which maps preserve positive semidefiniteness. Moreover, such maps act on an algebra and not only a vector space, and vector multiplication can play a role. Norms on  $\mathcal{B}(A)$  induce operator norms on  $\mathcal{B}(\mathcal{B}(A))$ , and  $\mathcal{B}(\mathcal{B}(A))$  can be equipped with its own norms as well.

We see that at each level properties are added but generally not removed. For example,  $\mathcal{B}(\mathcal{B}(A))$  can be seen as a vector space, as the set of operators on the vector space  $\mathcal{B}(A)$ , as the set of operators on the Hilbert space  $(\mathcal{B}(A), \langle \cdot, \cdot \rangle_{\text{HS}})$ , as the set of operators on the  $C^*$ -algebra  $(\mathcal{B}(A), \|\cdot\|_\infty)$ , as superoperators on a Hilbert space  $A$ , or even as a Hilbert space itself. Additionally, norms at lower levels induce operator norms at higher levels.

### 10.1.1 Vector spaces

The starting point is a vector space  $\mathcal{X}$  over  $\mathbb{R}$  or  $\mathbb{C}$ . Some examples:

- $A$  is a complex vector space of dimension  $d_A$
- $\mathcal{B}(A)$  is a complex vector space of (complex) dimension  $d_A^2$ , while  $\mathcal{B}_{\text{sa}}(A)$  is a real vector space of (real) dimension  $d_A^2$ .
- $\mathcal{B}(\mathcal{B}(A) \rightarrow \mathcal{B}(B))$  is a complex vector space of dimension  $d_A^2 d_B^2$ .

Vector spaces are closed under scalar multiplication and vector addition. Given two subsets  $S_1, S_2 \subseteq \mathcal{X}$ , we define

$$S_1 + S_2 = \{u + v : u \in S_1, v \in S_2\}.$$

A *subspace*  $V \subseteq \mathcal{X}$  is a subset which is closed under scalar multiplication and vector addition. If  $V_1, V_2 \subseteq \mathcal{X}$  are subspaces, then so is  $V_1 + V_2$  and  $V_1 \cap V_2$ . We say  $\mathcal{X}$  is the *direct sum* of subspaces  $V_1, \dots, V_n$  if

$$\mathcal{X} = V_1 + \dots + V_n$$

and  $V_i \cap V_j = \{0\}$  for each  $i \neq j$ . In that case, we write  $\mathcal{X} = V_1 \oplus V_2 \oplus \cdots \oplus V_n$ . Moreover, for any  $u \in \mathcal{X}$ , there exists a unique decomposition  $u = \sum_{i=1}^n v_i$  where  $v_i \in V_i$ ,  $i = 1, \dots, n$ . We can equip  $\mathcal{X}$  with a *norm* which is a function  $\|\cdot\| : \mathcal{X} \rightarrow \mathbb{R}_{\geq 0}$  which is subadditive (i.e. satisfies triangle inequality), positive definite ( $\|x\| = 0 \Rightarrow x = 0_{\mathcal{X}}$ ), and absolutely homogeneous (i.e.  $\|\alpha x\| = |\alpha| \|x\|$ ).

### 10.1.2 Operators on a vector space

Let  $\mathcal{X}$  be a vector space over  $\mathbb{C}$  with finite dimension  $d$ . Now, we consider  $\mathcal{B}(\mathcal{X})$ . Given a norm  $\|\cdot\|_{\mathcal{X}}$  on  $\mathcal{X}$ , the *operator norm* is defined by

$$\|R\|_{\mathcal{X} \rightarrow \mathcal{X}} := \sup_{\|x\|_{\mathcal{X}} \leq 1} \|Rx\|_{\mathcal{X}}$$

for  $R \in \mathcal{B}(\mathcal{X})$  is a norm on  $\mathcal{B}(\mathcal{X})$ . We denote the identity operator on  $\mathcal{X}$  as  $\mathbf{1}_{\mathcal{X}} \in \mathcal{B}(\mathcal{X})$ . Next, we say

1. An operator  $P \in \mathcal{B}(\mathcal{X})$  is a *projection* if  $P^2 = P$ . For any projection  $P$ , we have  $\mathcal{X} = P\mathcal{X} \oplus P^{\perp}\mathcal{X}$  where

$$P\mathcal{X} := \{Pu : u \in \mathcal{X}\}, \quad P^{\perp} := \mathbf{1}_{\mathcal{B}(\mathcal{X})} - P.$$

On the other hand, if  $\mathcal{X} = V_1 \oplus \cdots \oplus V_n$ , then  $P_i$  defined as  $P_i u = v_i$  where  $v_i$  is the unique element of  $V_i$  such that  $u = \sum_{i=1}^n v_i$  for  $v_i \in V_i$ , is a projection. Moreover,  $\sum_{i=1}^n P_i = \mathbf{1}_{\mathcal{X}} \in \mathcal{B}(\mathcal{X})$  is the identity operator, and  $P_i P_j = \delta_{ij} P_i$ .

Restating this, if  $\mathcal{X} = V_1 \oplus \cdots \oplus V_n$ , then any  $\Phi \in \mathcal{B}(\mathcal{X})$  can be decomposed as

$$\Phi = \sum_{ij} \Phi_{ij} \tag{10.3}$$

where  $\Phi_{ij} = P_j \Phi P_i$ , and the  $\Phi_{ij}$  commute with each other.

2. An operator  $N \in \mathcal{B}(\mathcal{X})$  is *nilpotent* if  $N^n = 0$  for some  $n \in \mathbb{N}$ . The *order* of a nilpotent operator  $N$  is the minimal  $n$  such that  $N^n = 0$ . The nilpotent order is always at most  $d$  and is denoted  $\text{ord } N$ .
3. An operator  $R \in \mathcal{B}(\mathcal{X})$  is a *contraction* on  $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$  if  $\|R\|_{\mathcal{X} \rightarrow \mathcal{X}} \leq 1$ .

Let  $\Phi \in \mathcal{B}(\mathcal{X})$  be a linear operator on  $\mathcal{X}$ . Then the *image*

$$\Phi(\mathcal{X}) := \{\Phi(x) : x \in \mathcal{X}\}$$



and *kernel*

$$\ker \Phi := \{x \in \mathcal{X} : \Phi(x) = 0\}$$

are subspaces of  $\mathcal{X}$ . The dimension of the image is called the rank of  $\Phi$ , denoted  $\text{rank}(\Phi)$ , and the Rank-Nullity Theorem states that  $\text{rank}(\Phi) + \dim \ker \Phi = d$ .

Now, let  $\mathcal{X}$  be a complex vector space of finite dimension  $d$ , and let  $\Phi \in \mathcal{B}(\mathcal{X})$ . Then  $\Phi$  has  $d$  (possibly repeated) *eigenvalues*  $\lambda \in \mathbb{C}$  and associated *eigenvectors*  $X \in \mathcal{X} \setminus \{0\}$  which are defined by the relationship

$$\Phi(X) = \lambda X.$$

The set  $\{X \in \mathcal{X} : \Phi(X) = \lambda X\}$  is a subspace of  $\mathcal{X}$  called the *eigenspace* of  $\Phi$  associated to the eigenvalue  $\lambda$ , and its dimension is the *geometric multiplicity* of  $\lambda$ .

Moreover,  $\Phi$  has a Jordan decomposition, defined as follows. Let  $n \leq d$  denote the number of distinct eigenvalues of  $\Phi$ , and let  $\mu_1, \dots, \mu_n$  denote the distinct eigenvalues of  $\Phi$ . Then there exists  $V_1, \dots, V_n$  such that

$$\mathcal{X} = V_1 \oplus \dots \oplus V_n \tag{10.4}$$

and each  $V_i$  is invariant under  $\Phi$ , namely  $\Phi(V_i) \subseteq V_i$  for each  $i = 1, \dots, n$ . The dimension of  $V_i$  is called the *algebraic multiplicity* of  $\mu_i$ , and is at least the geometric multiplicity.

Considering the decomposition of  $\Phi$  given in (10.3) induced by the direct sum decomposition of (10.4), we have  $\Phi_{ij} = 0$  for  $i \neq j$ , since  $\Phi P_i \subseteq P_i \mathcal{X}$  and hence  $P_j \Phi P_i \subseteq P_j P_i \mathcal{X} = \{0\}$ , where  $P_i$  is the projection onto  $V_i$ . Thus, it remains to understand the  $\Phi_{ii}$ . In fact, we have that  $\Phi_{ii} = \mu_i P_i + D_i$  where  $D_i$  is nilpotent. The projection  $P_i$  is called *spectral projection* of  $\Phi$  for eigenvalue  $\mu_i$ , and  $D_i$  the *eigen-nilpotent*. If  $\dim V_i = 1$ , then  $\mu_i$  is called a *simple* eigenvalue, and  $D_i$  is necessarily zero. If  $\dim V_i \geq 1$  and  $D_i = 0$ , then  $\mu_i$  is called a *semi-simple* eigenvalue. The resulting decomposition

$$\Phi = \sum_i \mu_i P_i + D_i \tag{10.5}$$

is known as the *Jordan decomposition* of  $\Phi$ . The *spectral radius* of a map  $\Phi \in \mathcal{B}(\mathcal{X})$  is defined by

$$\text{spr}(\Phi) := \max_i |\mu_i| \tag{10.6}$$

as the maximum modulus of its eigenvalues.

So far, we have investigated how the vector space structure of  $\mathcal{X}$  and the linearity of  $\Phi$  determines the structure of  $\Phi$  via its Jordan canonical form. Next, we show one way

in which norms on  $\mathcal{X}$  can be used to obtain information about the Jordan decomposition of  $\Phi$ .

**Lemma 10.1.1.** *Let  $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$  be a finite-dimensional Banach space (i.e. a complete normed vector space). The peripheral eigenvalues of a contraction  $\Phi \in \mathcal{B}(\mathcal{X})$  with spectral radius 1 on  $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$  are semi-simple.*

*Remark.* The following proof is reproduced from my master's thesis [Han16b, Lemma 3.1] for the sake of self-containedness; see also Prop. 6.2 of [Wol12].

*Proof.* Write  $\Phi$  in its Jordan canonical form,  $\Phi = \sum_i \mu_i P_i + D_i$ , where each  $\mu_i$  is an eigenvalue of  $\Phi$ ,  $P_i$  is the associated eigenprojection, and  $D_i$  the associated nilpotent, summed from  $i = 1$  to  $\dim \mathcal{X}$ . Let  $m_i = \text{ord } D_i$ , and  $m = \max\{m_i : 1 \leq i \leq \dim \mathcal{X}\}$ . Assume for some eigenvalue  $\mu_i$  with  $|\mu_i| = 1$  that  $m_i \geq 2$ , i.e.,  $\mu_i$  has an eigen-nilpotent. We wish to derive a contradiction, implying  $D_i \equiv 0$ .

Note that for  $n > m$ , using  $D_i P_j = P_j D_i = \delta_{ij} D_i$ , and  $P_i P_j = P_j P_i = \delta_{ij} P_i$ , a binomial expansion yields

$$\Phi^n = \sum_j \sum_{k=0}^m \binom{n}{k} \mu_j^{n-k} P_j (D_j)^k.$$

Our approach to deriving a contradiction will be to use that the binomial coefficient  $\binom{n}{k}$  becomes large with  $n$ , while  $\mu_i^n$  stays on the unit circle, and  $\|\Phi^n\| \leq \|\Phi\|^n = 1$  remains bounded.

Since  $D_i^{m_i-1} \neq 0$ , for some vector  $v$  we have  $D_i^{m_i-1} v \neq 0$ . Let  $n \in \mathbb{N}$  large enough so that  $n \|D_i^{m_i-1} v\|_{\mathcal{X}} > \|D_i^{m_i-2} v\|_{\mathcal{X}}$  and  $n > m$ . Then, since  $\Phi$  is a contraction,

$$\|D_i^{m_i-2} v\|_{\mathcal{X}} \geq \|\Phi^n D_i^{m_i-2} v\|_{\mathcal{X}} = \left\| \sum_j \sum_{k=0}^m \binom{n}{k} \mu_j^{n-k} P_j D_j^k D_i^{m_i-2} v \right\|_{\mathcal{X}}.$$

Note  $P_j P_i = 0$  implies  $P_j D_i^{m_i-2} v = P_j P_i D_i^{m_i-2} v = 0$  for  $i \neq j$ , so we have

$$\|D_i^{m_i-2} v\|_{\mathcal{X}} \geq \left\| \sum_{k=0}^m \binom{n}{k} \mu_i^{n-k} P_i (D_i)^k D_i^{m_i-2} v \right\|_{\mathcal{X}}.$$

But, by choice of  $m_i$ , only two terms in the sum survive:  $k = 0$  and  $k = 1$ :

$$\|D_i^{m_i-2} v\|_{\mathcal{X}} \geq \|\mu_i^n D_i^{m_i-2} v + n \mu_i^{n-1} D_i^{m_i-1} v\|_{\mathcal{X}}.$$

Then, by reverse triangle inequality and using  $|\mu_i| = 1$ ,

$$\|D_i^{m_i-2}v\|_{\mathcal{X}} \geq \left| \|D_i^{m_i-2}v\|_{\mathcal{X}} - n \|D_i^{m_i-1}v\|_{\mathcal{X}} \right| \geq n \|D_i^{m_i-1}v\|_{\mathcal{X}}.$$

By our choice of  $v$ , we have  $\|D_i^{m_i-1}v\|_{\mathcal{X}} \neq 0$ , hence

$$\frac{\|D_i^{m_i-2}v\|_{\mathcal{X}}}{\|D_i^{m_i-1}v\|_{\mathcal{X}}} \geq n$$

for all  $n$  large enough. This is a contradiction to our choice of  $n$ ; moreover, we could take  $n \rightarrow \infty$ , while the LHS remains bounded.  $\square$

### 10.1.3 Operators on a Hilbert space

Let  $\mathcal{X} = A$ , a Hilbert space with an inner product  $\langle \cdot, \cdot \rangle_A$ . For example:

- $\mathcal{H} = \mathbb{C}^{d_{\mathcal{H}}}$  for some  $d_{\mathcal{H}} \in \mathbb{N}$ , equipped with the inner product  $\langle u, v \rangle_{\mathcal{H}} := \sum_{i=1}^{d_{\mathcal{H}}} \bar{u}_i v_i$ .
- $\mathcal{B}(\mathcal{H})$ , for  $\mathcal{H}$  as in the previous example, equipped with Hilbert-Schmidt inner product (10.1).

The Hilbert space  $A$  has a privileged norm  $\|\psi\|_A := \sqrt{\langle \psi, \psi \rangle_A}$  for  $\psi \in A$ ; this norm satisfies the parallelogram law

$$2\|\psi\|_A^2 + 2\|\phi\|_A^2 = \|\psi + \phi\|_A^2 + \|\psi - \phi\|_A^2 \quad \forall \psi, \phi \in A.$$

The set of operators on  $\mathcal{B}(A)$  inherit the operator norm induced by  $\|\cdot\|_A$ , namely

$$\|X\|_{\infty} \equiv \|X\|_{\|\cdot\|_A \rightarrow \|\cdot\|_A} := \sup_{\substack{\psi \in A \\ \|\psi\|_A \leq 1}} \|A\psi\|_A. \quad (10.7)$$

The inner product on  $A$  also induces the adjoint  $*$  :  $\mathcal{B}(A) \rightarrow \mathcal{B}(A)$ , an anti-linear map defined via (10.2). The adjoint provides for a notion of self-adjointness and positive semidefiniteness. We say  $X \in \mathcal{B}(A)$  is *self-adjoint* if  $X = X^*$ , and is *positive semidefinite* if  $X = Y^*Y$  for some  $Y \in \mathcal{B}(A)$ , denoted  $X \geq 0$ . We see that if  $X \geq 0$  then

$$\langle \psi, X\psi \rangle_A = \langle \psi, Y^*Y\psi \rangle_A = \langle Y\psi, Y\psi \rangle_A = \|Y\psi\|_A^2 \geq 0 \quad \forall \psi \in A. \quad (10.8)$$

As mentioned at the start, the set of positive semidefinite operators on  $A$  are denoted  $\mathcal{B}_+(A)$ , and the set of full-rank positive semidefinite operators are denoted  $\mathcal{B}_{++}(A)$ . Note that if  $X \in \mathcal{B}_{++}(A)$ , then the inequality in (10.8) becomes strict unless  $\psi = 0_A$ .

The adjoint and the operator norm induced by the Hilbert space norm satisfy the so-called  $C^*$  identity,

$$\|X\|_\infty^2 = \|X^*\|_\infty^2 = \|X^*X\|_\infty \quad \forall X \in \mathcal{B}(A). \quad (10.9)$$

The operator norm is one member of an interpolating family of norms on  $\mathcal{B}(A)$ , the Schatten  $p$ -norms, defined as

$$\|X\|_p := (\text{tr}[|X|^p])^{1/p}$$

for  $p \geq 1$ ,  $X \in \mathcal{B}(A)$ , and  $|X| := \sqrt{X^*X}$ , where the function  $\sqrt{\cdot} : \mathcal{B}_+(A) \rightarrow \mathcal{B}_+(A)$  is defined as follows. Let  $X = \sum_i \mu_i P_i$  have a Jordan form (10.5) with no eigen-nilpotents (i.e.  $D_i = 0$  for all  $i$ ); this holds, for example, if  $X$  is self-adjoint. Then if  $f : D \subseteq \mathbb{C} \rightarrow \mathbb{C}$  is defined on a set  $D$  which includes the eigenvalues  $\{\mu_i\}_i$ , we define

$$f(X) := \sum_i f(\mu_i) P_i. \quad (10.10)$$

Since  $X^*X \geq 0$  for any  $X \in \mathcal{B}(A)$ , we have that the eigenvalues of  $X^*X$  are all non-negative, and hence the square-root  $\sqrt{X^*X}$  indeed can be defined in this manner.

As it turns out, the operator norm induced by the Hilbert space norm, (10.7), is exactly the limit  $p \rightarrow \infty$  of the Schatten  $p$ -norms, motivating the notation  $\|\cdot\|_\infty$ . Note also that the Schatten 1-norm is used to define the trace distance used extensively in Part I, and the Schatten 2-norm is the norm induced by the Hilbert-Schmidt inner product.

#### 10.1.4 Superoperators

A *superoperator* is a map  $\Phi \in \mathcal{B}(\mathcal{B}(A) \rightarrow \mathcal{B}(A'))$  for some Hilbert spaces  $A$  and  $A'$ . We denote the identity superoperator from  $\mathcal{B}(A) \rightarrow \mathcal{B}(A)$  as  $\text{id}_A$ . Note that  $\text{id}_A = \mathbf{1}_{\mathcal{B}(A)}$ . Given  $\Phi \in \mathcal{B}(\mathcal{B}(A) \rightarrow \mathcal{B}(A'))$ , we say

1.  $\Phi$  is *positive* (P) if  $\Phi(\mathcal{B}_+(A)) \subseteq \mathcal{B}_+(A')$ ;
2.  $\Phi$  is *hermitian-preserving* if  $\Phi(\mathcal{B}_{\text{sa}}(A)) \subseteq \mathcal{B}_{\text{sa}}(A')$ ;
3.  $\Phi$  is *completely positive* (CP) if  $\Phi \otimes \text{id}_E : \mathcal{B}(A \otimes E) \rightarrow \mathcal{B}(A' \otimes E)$  is positive, for any finite-dimensional system  $E$ ;
4.  $\Phi$  is *trace-preserving* (TP) if  $\text{tr}[\Phi(X)] = \text{tr}[X]$  for all  $X \in \mathcal{B}(A)$ ;
5.  $\Phi$  is *trace non-increasing* (TNI) if  $\text{tr}(\Phi(X)) \leq \text{tr}(X)$  for all  $X \in \mathcal{B}_+(A)$ ;

6.  $\Phi$  is *unital* (U) if  $\Phi(\mathbf{1}_A) = \mathbf{1}_{A'}$ ;
7.  $\Phi$  is a *quantum channel* if it is completely positive and trace-preserving (CPTP)
8. A quantum channel  $\Phi$  is *faithful* if it has a faithful invariant state (i.e.  $\sigma \in \mathcal{D}_+(A)$  such that  $\Phi(\sigma) = \sigma$ ).

Additionally,  $\Phi^*$  will denote the adjoint of  $\Phi$  with respect to the Hilbert-Schmidt inner product.

*Remark.*  $\Phi$  being hermitian-preserving is equivalent to the property that  $\Phi(X^*) = \Phi(X)^*$  for all  $X \in A$ , and is implied by positivity. The first statement follows from the fact that there is a unique decomposition  $X = \text{Re}(X) + i \cdot \text{Im}(X)$  such that  $\text{Re}(X), \text{Im}(X) \in \mathcal{B}_{\text{sa}}(A)$ , which is satisfied by  $\text{Re}(X) := \frac{X+X^*}{2}$  and  $\text{Im}(X) := \frac{X-X^*}{2i}$ . If  $\Phi$  is hermitian-preserving, then

$$\Phi(X)^* = [\Phi(\text{Re}(X)) + i\Phi(\text{Im}(X))]^* = \Phi(\text{Re}(X)) - i\Phi(\text{Im}(X)) = \Phi(X^*).$$

Conversely, if  $\Phi(X^*) = \Phi(X)^*$  for all  $X \in \mathcal{B}(A)$ , then for  $Y \in \mathcal{B}_{\text{sa}}(A)$ , we have

$$2i \text{Im}(\Phi(Y)) = 2i[\Phi(Y) - \Phi(Y)^*] = 2i[\Phi(Y - Y^*)] = 2i\Phi(0) = 0.$$

The second statement follows from the decomposition  $X = X_+ - X_-$  with  $X_{\pm} \in \mathcal{B}_+(A)$  which exists for  $X \in \mathcal{B}(A)$  if and only if  $X \in \mathcal{B}_{\text{sa}}(A)$ . Then if  $X \in \mathcal{B}_{\text{sa}}(A)$ , we have  $\Phi(X) = \Phi(X_+) - \Phi(X_-) \in \mathcal{B}_{\text{sa}}(A)$  as the difference of positive operators.

Note also that  $\Phi$  being positive implies that if  $X, Y \in \mathcal{B}_{\text{sa}}(A)$  satisfy  $X \leq Y$ , then  $\Phi(X) \leq \Phi(Y)$ . This follows immediately from the fact that  $Y - X \geq 0$ , so  $\Phi(Y - X) \geq 0$ .

One particular superoperator of relevance is that of conjugation. If  $R \in \mathcal{B}(A)$ , define  $\text{Ad}_R \in \mathcal{B}(\mathcal{B}(A) \rightarrow \mathcal{B}(A))$  by  $\text{Ad}_R(X) := RXR^*$  for  $X \in \mathcal{B}(A)$ . The notation for this operation is borrowed from [SS12].

The first result of interest concerns induced norms on  $\Phi$ . In the following,  $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$  represents a normed vector space.

**Theorem 10.1.2** (Russo Dye [RD66]). *Let  $\Phi : (\mathcal{B}(A), \|\cdot\|_{\infty}) \rightarrow (\mathcal{X}, \|\cdot\|_{\mathcal{X}})$  be a linear map. Then*

$$\|\Phi\|_{\|\cdot\|_{\infty} \rightarrow \|\cdot\|_{\mathcal{X}}} = \sup_{\substack{U \in \mathcal{B}(A) \\ U \text{ unitary}}} \|\Phi(U)\|_{\mathcal{X}}.$$

*If  $(\mathcal{X}, \|\cdot\|_{\mathcal{X}}) = (\mathcal{B}(B), \|\cdot\|_{\infty})$  and  $\Phi$  is unital, then  $\Phi$  is positive if and only if  $\|\Phi\|_{\|\cdot\|_{\infty} \rightarrow \|\cdot\|_{\infty}} = 1$ .*

The following is a well-known consequence of the Russo-Dye theorem: PTP linear maps are contractions on  $(\mathcal{B}(A), \|\cdot\|_1)$ .

**Corollary 10.1.3.** *Let  $\Phi : (\mathcal{B}(A), \|\cdot\|_1) \rightarrow (\mathcal{B}(A), \|\cdot\|_1)$  be positive and trace-preserving. Then*

$$\|\Phi\|_{\|\cdot\|_1 \rightarrow \|\cdot\|_1} = 1.$$

*Proof.* Let

$$\Phi' : (\mathcal{B}(A), \|\cdot\|_1)' \rightarrow (\mathcal{B}(A), \|\cdot\|_1)'$$

denote the *Banach space adjoint* of  $\Phi$  (also known as the topological dual), where  $(\mathcal{B}(A), \|\cdot\|_1)'$  is the set of continuous linear functionals on  $\mathcal{B}(A)$  equipped with the norm

$$\|\ell\|_1' = \sup_{x \in \mathcal{B}(A)} |\ell(x)|$$

and where  $\Phi'$  is defined by  $\Phi'(\ell)(x) = \ell(\Phi(x))$ . In particular,  $\Phi'(\text{tr})(x) = \text{tr}(\Phi(x)) = \text{tr}(x)$ , so  $\Phi'(\text{tr}) = \text{tr}$ .

In fact, the Banach dual of  $(\mathcal{B}(A), \|\cdot\|_1)$  is isomorphic to  $(\mathcal{B}(A), \|\cdot\|_\infty)$  (see [RS81, Theorem VI.26]) and each  $\ell \in (\mathcal{B}(A), \|\cdot\|_1)'$  is of the form

$$\ell(x) = \text{tr}[y_\ell^* x] = \langle y_\ell, x \rangle_{\text{HS}}$$

for some  $y_\ell \in \mathcal{B}(A)$ , and of course each  $y_\ell \in \mathcal{B}(A)$  gives such a linear functional. Under the map  $\iota : \ell \mapsto y_\ell$ , we see that  $\text{tr} \mapsto \mathbf{1}_A$ . Additionally,

$$\|\ell\|_1' = \sup_{x \in \mathcal{B}(A)} |\ell(x)| = \sup_{x \in \mathcal{B}(A)} |\text{tr}[y_\ell x]| = \|y_\ell\|_\infty.$$

Thus,  $\iota$  is an anti-linear isometry from  $(\mathcal{B}(A), \|\cdot\|_1)'$  to  $(\mathcal{B}(A), \|\cdot\|_\infty)$ . Moreover,  $\tilde{\Phi} = \iota \circ \Phi' \circ \iota^{-1}$  satisfies

$$\tilde{\Phi}(y_\ell) = \iota \circ \Phi' \circ \iota^{-1} \circ \iota(\ell) = \iota \circ \Phi'(\ell) = y_{\Phi'(\ell)}.$$

Hence,

$$\langle \tilde{\Phi}(y_\ell), x \rangle_{\text{HS}} = \langle y_{\Phi'(\ell)}, x \rangle_{\text{HS}}.$$

Then expanding the right-hand side,

$$\langle y_{\Phi'(\ell)}, x \rangle_{\text{HS}} = \Phi'(\ell)(x) = \ell(\Phi(x)) = \langle y_\ell, \Phi(x) \rangle_{\text{HS}}.$$

Thus, we've found that  $\tilde{\Phi}$  is in fact  $\Phi^*$ , the Hilbert-Schmidt adjoint. Since Hilbert-Schmidt adjoints preserve positivity, we see that  $\tilde{\Phi}$  is positive. Thus, by the Russo-Dye theorem,

$$\|\tilde{\Phi}\|_{\infty \rightarrow \infty} = \|\Phi^*\|_{\infty \rightarrow \infty} = 1.$$

Moreover,

$$\begin{aligned} \|\tilde{\Phi}\|_{\infty \rightarrow \infty} &= \sup_{\|x\|_\infty=1} \|\tilde{\Phi}(x)\|_\infty = \sup_{\|\ell\|'_1=1} \|\tilde{\Phi}(\iota(\ell))\|_\infty \\ &= \sup_{\|\ell\|'_1=1} \|\iota \circ \Phi'(\ell)\|_\infty = \sup_{\|\ell\|'_1=1} \|\Phi'(\ell)\|'_1 = \|\Phi'\|_{\|\cdot\|'_1 \rightarrow \|\cdot\|'_1}. \end{aligned}$$

But since the Banach space adjoint is an isometry, we have  $\|\Phi'\|_{\|\cdot\|'_1 \rightarrow \|\cdot\|'_1} = \|\Phi\|_{\|\cdot\|_1 \rightarrow \|\cdot\|_1}$ , concluding the proof.  $\square$

We also find that the spectral radius of a PTP linear map  $\Phi$  is 1. Clearly, the spectral radius cannot be greater than one; if  $\Phi(X) = \lambda X$  for  $|\lambda| > 1$ , then  $\|\Phi(X)\|_1 = |\lambda| \|X\|_1$  and hence  $\|\Phi\|_{\|\cdot\|_1 \rightarrow \|\cdot\|_1} \geq |\lambda| > 1$ , contradicting Corollary 10.1.3. On the other hand, since  $\Phi^*$  is unital we have  $1 \in \text{spec}(\Phi^*) = \overline{\text{spec}(\Phi)}$  and hence  $1 \in \text{spec}(\Phi)$  as well. Thus,  $\text{spr}(\Phi) = 1$ , and in fact, 1 itself is an eigenvalue of  $\Phi$ . These properties and Lemma 10.1.1 show that the peripheral eigenvalues of a PTP linear map are semi-simple, meaning their associated eigen-nilpotents vanish. We can also check that the spectrum is closed under complex conjugation, as follows. If  $\Phi$  is positive, then it is Hermitian preserving. Hence  $\Phi(X) = \lambda X$  implies  $\Phi(X^*) = \Phi(X)^* = (\lambda X)^* = \bar{\lambda} X^*$ .

We have started to build an understanding of the spectral properties of PTP linear maps: they are contractions in 1-norm, 1 is the spectral radius and an eigenvalue, all peripheral eigenvalues are semi-simple, and the spectrum is closed under complex conjugation. We will return to this study of the spectral properties of these maps, in the completely positive case, in Section 10.3. First, we will review the notions of separability and entanglement.

## 10.2 Separability and entanglement

In this section, we define separable operators and correspondingly, entanglement-breaking superoperators, and investigate some of their basic properties and relationships.

### 10.2.1 Separability at the level of operators

**Definition 10.2.1** ( $\text{SEP}(A : B)$ ). We define the cone of separable operators in  $\mathcal{B}(A \otimes B)$  with respect to the partition  $A : B$  as

$$\text{SEP}(A : B) := \text{conv} \{X_A \otimes Y_B : X_A \in \mathcal{B}_+(A), Y_B \in \mathcal{B}_+(B)\} \subseteq \mathcal{B}_+(A \otimes B).$$

**Definition 10.2.2** (Entanglement witness). Let  $X_{AB} \in \mathcal{B}_+(A \otimes B)$ . An *entanglement witness*  $W_{AB} \in \mathcal{B}_{\text{sa}}(A \otimes B)$  for  $X_{AB}$  is any operator such that

1.  $\langle W_{AB}, X_{AB} \rangle_{\text{HS}} < 0$
2.  $\langle W_{AB}, Z_{AB} \rangle_{\text{HS}} \geq 0$  for all  $Z_{AB} \in \text{SEP}(A : B)$ .

As the name suggests, entanglement witnesses provide a “witness” (or “certificate”) that a given operator is entangled. We can see directly from the definition that an entanglement witness for  $X_{AB}$  cannot exist if  $X_{AB} \in \text{SEP}(A : B)$ . The following proposition shows that for any entangled  $X_{AB}$ , an entanglement witness always exists. This is a consequence of the fact that  $\text{SEP}(A : B)$  is a closed convex set, and if  $X \notin \text{SEP}(A : B)$ , then there exists a hyperplane separating  $X_{AB}$  from  $\text{SEP}(A : B)$ .

**Proposition 10.2.3** ([HHH96]).  $X_{AB} \in \mathcal{B}_+(A \otimes B)$  has an entanglement witness if and only if  $X_{AB} \notin \text{SEP}(A : B)$ .

The following is a simple but very useful lemma regarding the structure of  $\text{SEP}$ . Roughly, it states that if the sum of two positive operators is separable, and locally they live in orthogonal subspaces, then each is separable.

**Lemma 10.2.4.** If  $X_{AB}, Y_{AB} \in \mathcal{B}_+(A \otimes B)$  satisfy

$$X_{AB} + Y_{AB} \in \text{SEP}(A : B)$$

and are such that for some orthogonal projection  $P \in \mathcal{B}(A)$  we have  $X_{AB} \in \mathcal{B}_+(PA \otimes B)$  and  $Y_{AB} \in \mathcal{B}_+(P^\perp A \otimes B)$ , then  $X_{AB} \in \text{SEP}(A : B)$  and  $Y_{AB} \in \text{SEP}(A : B)$ . Here,  $P^\perp = \mathbf{1}_A - P$ , and e.g.

$$PA \otimes B := \{(P \otimes \mathbf{1}_B) |\psi_{AB}\rangle : |\psi_{AB}\rangle \in A \otimes B\} \subseteq A \otimes B.$$

*Proof.* Assume  $X_{AB} \notin \text{SEP}(A : B)$ . Then there exists an entanglement witness  $W_{AB}$  for  $X_{AB}$  by Definition 10.2.2. Let  $\tilde{W}_{AB} = \text{Ad}_P \otimes \text{id}(W_{AB})$ ; recall  $\text{Ad}_P$  denotes the



superoperator  $X \mapsto PXP^*$ . Then  $\tilde{W}_{AB}$  is an entanglement witness for  $X_{AB} + Y_{AB}$ :

$$\mathrm{tr}[\tilde{W}_{AB}(X_{AB} + Y_{AB})] = \mathrm{tr}[W_{AB}(X_{AB} + \mathrm{Ad}_P \otimes \mathrm{id}(Y_{AB}))] = \mathrm{tr}[W_{AB}X_{AB}] < 0$$

since  $\mathrm{Ad}_P^* = \mathrm{Ad}_{P^*} = \mathrm{Ad}_P$  and  $P(I - P) = 0$ , and similarly for  $Z_{AB} \in \mathrm{SEP}(A : B)$ ,

$$\mathrm{tr}[\tilde{W}_{AB}Z_{AB}] = \mathrm{tr}[W_{AB} \mathrm{Ad}_P \otimes \mathrm{id}(Z_{AB})] \geq 0$$

since  $\mathrm{Ad}_P \otimes \mathrm{id}(Z_{AB}) \in \mathrm{SEP}(A : B)$  by Lemma 10.2.12.  $\square$

### 10.2.2 Choi-Jamiolkowski (CJ) isomorphism: translate between operators and superoperators

**Definition 10.2.5.** A linear map  $\Phi : \mathcal{B}(A) \rightarrow \mathcal{B}(A')$  is *entanglement-breaking* (EB) if  $\Phi \otimes \mathrm{id}_B(\mathcal{B}_+(A \otimes B)) \subseteq \mathrm{SEP}(A : B)$  for all  $B$ .

In fact, it suffices to take the dimension of the  $B$  system equal to  $d_A$ . Note also that (EB) implies (CP).

**Definition 10.2.6** (CJ).  $J_b(\Phi) := (\Phi \otimes \mathrm{id}_{\tilde{A}})(\Omega_{A:\tilde{A}}^b) \in \mathcal{B}(A' \otimes \tilde{A})$  where  $\tilde{A}$  is a copy of  $A$ , and

$$|\Omega_{A:B}^b\rangle := \sum_{i=1}^{\min(d_A, d_B)} |i_A\rangle \otimes |i_B\rangle, \quad \Omega_{A:B}^b := |\Omega_{A:B}^b\rangle \langle \Omega_{A:B}^b|$$

and  $\{|i_A\rangle\}_{i=1}^{d_A}$  and  $\{|i_B\rangle\}_{i=1}^{d_B}$  are orthonormal bases (ONB) of  $A$  and  $B$  respectively, and  $b = (\{|i_A\rangle\}_{i=1}^{d_A}, \{|i_B\rangle\}_{i=1}^{d_B})$  denotes the pair of ONBs. We usually regard  $J_b(\Phi) \in \mathcal{B}(A' \otimes A)$  for notational simplicity, and write  $J \equiv J_b$  when the choice of basis does not matter (which is most of the time), and likewise  $\Omega_{A:B} \equiv \Omega_{A:B}^b$ .

**Proposition 10.2.7.**  $J_b : (\mathcal{B}(\mathcal{B}(A) \rightarrow \mathcal{B}(A')), \langle \cdot, \cdot \rangle_{\mathrm{HS}}) \rightarrow (\mathcal{B}(A' \otimes A), \langle \cdot, \cdot \rangle_{\mathrm{HS}})$  is a linear bijective isometry, i.e. a Hilbert space isomorphism, and the value  $\|J_b(\Phi)\|_{\mathrm{HS}}$  is independent of the choice of bases  $b$ .

*Proof.* Again letting  $\tilde{A}$  be a copy of  $A$  for explicitness, we have

$$\begin{aligned} \|J_b(\Phi)\|_{\mathcal{B}(A' \otimes \tilde{A}), \mathrm{HS}}^2 &= \langle (\Phi \otimes \mathrm{id})(\Omega_{A:\tilde{A}}^b), (\Phi \otimes \mathrm{id})(\Omega_{A:\tilde{A}}^b) \rangle_{\mathcal{B}(A' \otimes \tilde{A}), \mathrm{HS}} \\ &= \langle \Omega_{A:\tilde{A}}^b, (\Phi^* \Phi \otimes \mathrm{id})(\Omega_{A:\tilde{A}}^b) \rangle_{\mathcal{B}(A \otimes \tilde{A}), \mathrm{HS}} \\ &= \sum_{i,j,i',j'=1}^{d_A} \langle |i_A\rangle \langle j_A| \otimes |i_{\tilde{A}}\rangle \langle j_{\tilde{A}}|, (\Phi^* \Phi \otimes \mathrm{id})(|i'_A\rangle \langle j'_A| \otimes |i'_{\tilde{A}}\rangle \langle j'_{\tilde{A}}|) \rangle_{\mathcal{B}(A \otimes \tilde{A}), \mathrm{HS}} \end{aligned}$$

$$\begin{aligned}
&= \sum_{i,j,i',j'=1}^{d_A} \langle i_A | \langle j_A |, \Phi^* \Phi(|i'_A\rangle \langle j'_A|) \rangle_{\mathcal{B}(A), \text{HS}} \underbrace{\langle i_{\bar{A}} | \langle j_{\bar{A}} |, |i'_{\bar{A}}\rangle \langle j'_{\bar{A}}| \rangle_{\mathcal{B}(\bar{A}), \text{HS}}}_{\delta_{ii'} \delta_{jj'}} \\
&= \sum_{i,j=1}^{d_A} \langle i_A | \langle j_A |, \Phi^* \Phi(|i_A\rangle \langle j_A|) \rangle_{\mathcal{B}(A), \text{HS}}
\end{aligned}$$

Since  $\{|i_A\rangle \langle j_A|\}_{i,j=1}^{d_A}$  is an ONB of  $\mathcal{B}(A)$ , then

$$\sum_{i,j=1}^{d_A} \langle i_A | \langle j_A |, \Phi^* \Phi(|i_A\rangle \langle j_A|) \rangle_{\mathcal{B}(A), \text{HS}} = \text{tr}[\Phi^* \Phi] = \langle \Phi, \Phi \rangle_{\mathcal{B}(\mathcal{B}(A) \rightarrow \mathcal{B}(A')), \text{HS}}$$

completing the proof that  $J_b$  is an isometry. Note also that the value  $\|J_b(\Phi)\|_{\text{HS}}$  does not depend on the choice of basis  $b$  since the trace is independent of basis.  $\square$

**Definition 10.2.8.** If  $\Phi : \mathcal{B}(A) \rightarrow \mathcal{B}(A')$  is a linear map, we say  $\{R_1, \dots, R_n\}$  with  $R_j \in \mathcal{B}(A \rightarrow A')$  is a Kraus decomposition for  $\Phi$  if  $\Phi = \sum_{i=1}^n \text{Ad}_{R_i}$ .

The following is a standard result which provides useful characterizations of complete positivity. The proof is omitted for the sake of brevity.

**Proposition 10.2.9.**  $\Phi \in \mathcal{B}(\mathcal{B}(A) \rightarrow \mathcal{B}(A'))$  has a Kraus decomposition if and only if it is CP if and only if  $J(\Phi) \geq 0$ .

The maximally-entangled operator  $\Omega_{A:B}$  satisfies some properties that will be helpful in the following: the “reflection trick” (also known as the “ricochet property”), and that any pure state  $|\psi\rangle_{AB}$  can be obtained from  $|\Omega\rangle_{A:B}$  via a local operation.

**Lemma 10.2.10** (Reflection trick). *For any  $A$  and  $B$  and  $R \in \mathcal{B}(B)$*

$$(\mathbf{1}_A \otimes R) |\Omega_{A:B}\rangle = (R^{T_b} \otimes \mathbf{1}_B) |\Omega_{A:B}\rangle$$

where  $R^{T_b} \in \mathcal{B}(A)$  is a “pseudo-transpose” (in the sense of the Moore-Penrose pseudo-inverse) of  $R$  induced by the pair of bases  $b$ : if  $R = \sum_{i,j=1}^{d_B} R_{ij} |i\rangle_B \langle j|_B$ , then

$$R^{T_b} = \sum_{i,j=1}^{\min(d_A, d_B)} R_{ji} |i\rangle_A \langle j|_A \in \mathcal{B}(A).$$

Likewise, if  $S \in \mathcal{B}(A)$ , then

$$(S \otimes \mathbf{1}_B) |\Omega_{A:B}\rangle = (\mathbf{1}_A \otimes S^{T_b}) |\Omega_{A:B}\rangle$$

where for  $S = \sum_{i,j=1}^{d_A} S_{ij} |i\rangle_A \langle j|_A$

$$S^{T_b} = \sum_{i,j=1}^{\min(d_A, d_B)} S_{ji} |i\rangle_B \langle j|_B \in \mathcal{B}(B)$$

**Lemma 10.2.11.** *For any  $A$  and  $B$  and pure state  $|\phi_{AB}\rangle \in A \otimes B$ , there exist  $R_1 \in \mathcal{B}(A)$  and  $R_2 \in \mathcal{B}(B)$  such that*

$$|\phi_{AB}\rangle = (R_1 \otimes \mathbf{1}_B) |\Omega_{A:B}\rangle = (\mathbf{1}_A \otimes R_2) |\Omega_{A:B}\rangle$$

where  $\text{tr}[R_1^\dagger R_1] = \text{tr}[R_2^\dagger R_2] = 1$ .

*Proof.* If  $|\phi_{AB}\rangle \in A \otimes B$ , then the Schmidt decomposition yields

$$|\phi_{AB}\rangle = \sum_{i=1}^{\min(d_A, d_B)} \sqrt{\lambda_i} |e_i\rangle \otimes |f_i\rangle$$

where  $\sum_{i=1}^{\min(d_A, d_B)} \lambda_i = \text{tr} \phi_{AB} = \langle \phi_{AB} | \phi_{AB} \rangle$ , and  $|e_i\rangle_{i=1}^{d_A}$  is an ONB of  $A$ , and  $|f_i\rangle_{i=1}^{d_B}$  is an ONB of  $B$ . Then

$$|\phi_{AB}\rangle = (R_A \otimes U_B) |\Omega_{A:B}\rangle$$

where we define  $U_B |i\rangle = |f_i\rangle$  for  $i = 1, \dots, d_B$ , and  $R_A |i\rangle = \sqrt{\lambda_i} |e_i\rangle$  for  $i = 1, \dots, d_A$ , where  $\lambda_i = 0$  for  $i > \min(d_A, d_B)$ . Then by Lemma 10.2.10,  $R_1 := R_A U_B^{T_b}$  suffices. We see also  $\text{tr}[R_1^\dagger R_1] = \text{tr}[R_A^\dagger R_A] = \sum_{i=1}^{d_A} \lambda_i = 1$ .  $R_2$  can be constructed similarly.  $\square$

### 10.2.3 EB maps: separability at the level of superoperators

**Lemma 10.2.12** (Local maps cannot create entanglement). *For any positive maps  $\Phi_1 : \mathcal{B}(A) \rightarrow \mathcal{B}(A')$  and  $\Phi_2 : \mathcal{B}(B) \rightarrow \mathcal{B}(B')$ ,*

$$(\Phi_1 \otimes \Phi_2)(\text{SEP}(A : B)) \subseteq \text{SEP}(A' : B').$$

*Proof.* If  $X_{AB} \in \text{SEP}(A : B)$ , then  $X_{AB} = \sum_{i=1}^n \lambda_i x_i \otimes y_i$  for  $n \in \mathbb{N}$ ,  $x_i \in \mathcal{B}_+(A)$ ,  $y_i \in \mathcal{B}_+(B)$  and  $\lambda_i \geq 0$ . Then

$$(\Phi_1 \otimes \Phi_2)(X_{AB}) = \sum_{i=1}^n \lambda_i \Phi_1(x_i) \otimes \Phi_2(y_i) \in \text{SEP}(A' : B')$$

since  $\Phi_1(x_i) \in \mathcal{B}_+(A')$  and  $\Phi_2(y_i) \in \mathcal{B}_+(B')$ .  $\square$

*Remark.* In particular, this means that if  $\Phi$  is entanglement-breaking and  $\Psi$  is positive, then  $\Psi \circ \Phi$  is entanglement-breaking, since  $J(\Psi \circ \Phi) = (\Psi \otimes \text{id})(J(\Phi))$ , and hence  $J(\Psi \circ \Phi)$  is the result of local positive operations applied to a separable state. Additionally, if  $\Psi$  is CP, then  $\Phi \circ \Psi$  is entanglement-breaking too, since  $J(\Phi \circ \Psi) = (\Phi \otimes \text{id})J(\Psi) \in \text{SEP}$  as  $J(\Psi) \geq 0$ .

**Corollary 10.2.13.** *If  $\Phi : \mathcal{B}(A) \rightarrow \mathcal{B}(A')$  is a linear map such that  $J_b(\Phi) \in \text{SEP}(A' : A)$  for some pair of bases  $b$ , then  $J_{b'}(\Phi) \in \text{SEP}(A' : B)$  for any other choice of pairs of bases  $b'$ .*

*Proof.*  $J_{b'}(\Phi) = \text{Ad}_{U_A} \otimes \text{Ad}_{U_B}(J_b(\Phi))$  for some unitaries  $U_A \in \mathcal{B}(A)$  and  $U_B \in \mathcal{B}(B)$ .  $\square$

We may use CJ to lift Lemma 10.2.4 to the superoperator level.

**Corollary 10.2.14** (Corollary to Lemma 10.2.4). *If  $\Phi_1$  and  $\Phi_2$  are CP maps from  $\mathcal{B}(A)$  to  $\mathcal{B}(A')$  such that their images are orthogonal, then  $J(\Phi_1 + \Phi_2) \in \text{SEP}(A' : A)$  if and only if both  $J(\Phi_1) \in \text{SEP}(A' : A)$  and  $J(\Phi_2) \in \text{SEP}(A' : A)$ .*

*In particular,  $J(\Phi_1 \oplus \Phi_2) \in \text{SEP}(A' : A)$  if and only if  $J(\Phi_1), J(\Phi_2) \in \text{SEP}(A' : A)$ .*

*Proof.* Let  $P \in \mathcal{B}(A')$  be the projection onto the image of  $\Phi_1$ . Then  $J(\Phi_1) = \text{Ad}_P \otimes \text{id}(J(\Phi_1))$  and  $J(\Phi_2) = \text{Ad}_{P^\perp} \otimes \text{id}(J(\Phi_2))$  for  $P^\perp = I - P$ . The result follows from Lemma 10.2.4.  $\square$

The following proposition shows that the CJ isomorphism preserves notions of separability between states and maps.

**Proposition 10.2.15.** *Let  $\Phi : \mathcal{B}(A) \rightarrow \mathcal{B}(A')$  be a CP map. Then  $\Phi$  is EB if and only if  $J(\Phi) \in \text{SEP}(A' : A)$ .*

*Remark.* [HSR03] proves this in the TP case. The fact is known in the general case, and I include the proof for completeness.

*Proof.* The implication  $\Rightarrow$  is immediate. For the reverse implication, let  $|\phi_{AB}\rangle \in A \otimes B$ . Then  $\phi_{AB} = \text{id}_A \otimes \text{Ad}_R(\Omega_{A:B})$  for some  $R \in \mathcal{B}(B)$  by Lemma 10.2.11. Then

$$\Phi \otimes \text{id}_B(\phi_{AB}) = \Phi \otimes \text{Ad}_R(\Omega_{A:B}) = \text{id} \otimes \text{Ad}_R(\tilde{J})$$

for

$$\tilde{J} := \Phi \otimes \text{id}_B(\Omega_{A:B}) = \sum_{i,j=1}^{\min(d_A, d_B)} \Phi(|i\rangle\langle j|) \otimes |i\rangle\langle j|.$$

By Lemma 10.2.12, it suffices to prove that  $\tilde{J} \in \text{SEP}(A : B)$ . Let us compare to

$$J(\Phi) = \sum_{i,j=1}^{d_A} \Phi(|i\rangle\langle j|) \otimes |i\rangle\langle j| \in \text{SEP}(A' : A)$$

- If  $d_A > d_B$ , then we can regard  $B = \text{span}\{|i\rangle\}_{i=1}^{d_B}$  as a subspace of  $A = \text{span}\{|i\rangle\}_{i=1}^{d_A}$ . Letting  $P$  denote the orthogonal projection onto this subspace (i.e.  $P = \sum_{i=1}^{d_B} |i\rangle\langle i| \in \mathcal{B}(A)$ ), then  $\tilde{J} = \text{id}_{A'} \otimes \text{Ad}_P(J(\Phi))$  and hence is separable by Lemma 10.2.12.
- If  $d_A \leq d_B$  then in the same way, we can regard  $A$  as a subspace of  $B$ , and note that  $\tilde{J}$  is in fact an element of  $\mathcal{B}(A' \otimes A) \subseteq \mathcal{B}(A' \otimes B)$ . Moreover,  $\tilde{J} = J(\Phi) \in \text{SEP}(A' : A)$ , concluding the proof. □

*Remark.* Proposition 10.2.15 implies that Corollary 10.2.14 directly yields statements about entanglement-breakingness.

## 10.3 Peripheral spectral properties of superoperators

In this section we consider the peripheral spectral properties of superoperators. Note that in this section, we mostly restrict to the case of endomorphisms in order to study spectral properties, which require the input and output space to be the same. The peripheral spectrum, i.e. the set of eigenvalues of maximal modulus, of a map  $\Phi \in \mathcal{B}(\mathcal{B}(A))$  plays a fundamental role in understanding the action of  $\Phi^n$  for large  $n$ . This arises in the notion of an *eventually* entanglement-breaking map. A map  $\Phi \in \mathcal{B}(\mathcal{B}(A))$  is *eventually-entanglement breaking* (EEB) if the  $n$ -fold composition  $\Phi^n$  is entanglement-breaking for some  $n \in \mathbb{N}$ . If  $\Phi, \Psi \in \mathcal{B}(\mathcal{B}(A))$ , and either is entanglement-breaking, then  $\Phi \circ \Psi$  is entanglement-breaking too. Thus, if  $\Phi^n$  is EB, so is  $\Phi^{n'}$  for all  $n' \geq n$ . Therefore, to know whether or not  $\Phi$  is EEB, it suffices to check for large  $n$ .

It is then natural to introduce the *phase subspace*  $\tilde{\mathcal{N}}(\Phi)$ , which is the union of the eigenspaces associated to peripheral eigenvalues. It is defined as

$$\tilde{\mathcal{N}}(\Phi) = \text{span}\{X \in \mathcal{B}(\mathcal{H}) : \exists \phi \in \mathbb{R} \text{ s.t. } \Phi(X) = e^{i\phi} X\},$$

i.e. the linear span of the peripheral points and denote by  $P$  the projection onto it. First, it's important to note that the phase subspace of a quantum channel always contains a density matrix. The following is proven in [Wol12, Theorem 6.11] as an immediate corollary to Brouwer's fixed point theorem.

**Proposition 10.3.1** (Existence of an invariant state). *Let  $\Phi \in \mathcal{B}(\mathcal{B}(A))$  be positive and trace-preserving. Then  $\Phi$  admits an invariant state  $\sigma \in \mathcal{D}(A)$ , i.e. such that  $\Phi(\sigma) = \sigma$ .*

*Remark.* By a very different proof technique, this can be generalized to remove the trace-preserving assumption; see [EH78, Theorem 2.5]. In this case, the spectral radius of  $\Phi$  may not be 1, and one obtains  $\sigma \geq 0$  such that  $\Phi(\sigma) = \text{spr}(\Phi)\sigma$ .

The invariant state  $\sigma$  guaranteed to exist by Proposition 10.3.1 may not be full-rank. Recall that if a quantum channel  $\Phi$  admits a full-rank invariant state, it is called *faithful*.

Moreover, for any quantum channel  $\Phi : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H})$ , the phase subspace is known to possess the following structure (Theorem 6.16 of [Wol12], Theorem 8 of [WP10]): there exists a decomposition of  $\mathcal{H}$  as  $\mathcal{H} = \bigoplus_{j=1}^K \mathcal{H}_j \otimes \mathcal{K}_j \oplus \mathcal{K}_0$  such that

$$\tilde{\mathcal{N}}(\Phi) := \bigoplus_{i=1}^K \mathcal{B}(\mathcal{H}_i) \otimes \tau_i \oplus 0_{\mathcal{K}_0}, \quad P(\rho \oplus 0_{\mathcal{K}_0}) = \sum_{i=1}^K \text{tr}_{\mathcal{K}_i}(p_i \rho p_i) \otimes \tau_i, \quad (10.11)$$

where  $p_i$  is the orthogonal projector onto the  $i$ -th subspace, for some fixed states  $\tau_i \in \mathcal{D}_+(\mathcal{K}_i)$ . Moreover, there exist unitaries  $U_i \in \mathcal{H}_i$ , and a permutation  $\pi \in S_K$  such that for any element  $X \in \tilde{\mathcal{N}}(\Phi)$ , we have

$$\Phi(X) = \bigoplus_{i=1}^K U_i X_{\pi(i)} U_i^\dagger \otimes \tau_i \oplus 0_{\mathcal{K}_0} \quad (10.12)$$

where  $X$  is decomposed as  $X = \bigoplus_{i=1}^K X_i \otimes \tau_i \oplus 0_{\mathcal{K}_0}$  according to (10.11). The permutation  $\pi$  permutes within subsets of  $\{1, \dots, K\}$  for which the corresponding  $\mathcal{H}_i$ 's have equal dimension. Note that the space  $\mathcal{K}_0 = \{0\}$  if and only if the quantum channel is faithful.

To further explore the spectral properties of  $\Phi \in \mathcal{B}(\mathcal{B}(A))$ , we need to introduce the notions of primitivity and irreducibility.

**Definition 10.3.2** (Primitive map). Let  $\Phi \in \mathcal{B}(\mathcal{B}(A))$  be positive and trace-preserving. Then we say  $\Phi$  is *primitive* if there exists a full-rank state  $\sigma \in \mathcal{D}(A)$  such that we have the long-time convergence

$$\Phi^n(\rho) \xrightarrow{n \rightarrow \infty} \sigma, \quad \forall \rho \in \mathcal{D}(A). \quad (10.13)$$

*Remark.* The limit (10.13) implies

$$\Phi^n(X) \xrightarrow{n \rightarrow \infty} \text{tr}[X]\sigma, \quad \forall X \in \mathcal{B}(A). \quad (10.14)$$

To see this, let  $X \in \mathcal{B}(A)$  and consider the decomposition  $X = X_+ - X_-$  where  $X_{\pm} \geq 0$ . Then  $\rho_{\pm} = \frac{X_{\pm}}{\text{tr}[X_{\pm}]} \in \mathcal{D}(A)$ , so  $\lim_{n \rightarrow \infty} \Phi^n(\rho_{\pm}) = \sigma$ . Hence, by linearity of  $\Phi$ ,  $X = \text{tr}[X_+]\rho_+ - \text{tr}[X_-]\rho_-$  satisfies

$$\lim_{n \rightarrow \infty} \Phi^n(X) = \text{tr}[X_+]\sigma - \text{tr}[X_-]\sigma = \text{tr}[X]\sigma$$

as desired.

**Proposition 10.3.3** (Spectral characterization of primitivity). *A positive trace-preserving map  $\Phi \in \mathcal{B}(\mathcal{B}(A))$  is primitive if and only if 1 is a non-degenerate eigenvalue of  $\Phi$ , 1 is the only eigenvalue of  $\Phi$  on the unit circle, and the eigenvector associated to 1 can be chosen to be strictly positive definite. Moreover, if  $\Phi$  is primitive, then the state  $\sigma$  in (10.13) is the strictly positive-definite eigenvector of  $\Phi$  associated to the eigenvalue 1.*

*Remark.* The definition of primitivity given in Definition 10.3.2 is a priori difficult to check, while the characterization in Proposition 10.3.3 can be checked in time polynomial in  $d_A$  by computing the peripheral eigenvalues of  $\Phi$  and possibly the eigenvalues of the eigenvector associated to the eigenvalue 1 (if it is indeed an eigenvalue).

*Proof.* Assume  $\Phi$  is primitive. Then if  $\Phi$  has an eigenvalue  $\lambda \in \mathbb{C}$  and corresponding eigenvector  $X \in \mathcal{B}(A)$  with  $|\lambda| = 1$ , we have  $\Phi^n(X) = \lambda^n X$  and thus

$$\text{tr}[X]\sigma = \lim_{n \rightarrow \infty} \Phi^n(X) = \lim_{n \rightarrow \infty} \lambda^n X = X \lim_{n \rightarrow \infty} \lambda^n.$$

Since  $|\lambda| = 1$  and  $X$  is an eigenvector, the right-hand side must not be zero. Moreover, for the limit to exist, we must have  $\lambda = 1$ ; this follows from the fact that

$$\lim_{n \rightarrow \infty} \lambda^n = \lim_{n \rightarrow \infty} \lambda^{n+1} = \lambda \lim_{n \rightarrow \infty} \lambda^n$$

and that  $\lambda \neq 0$ . Thus, the only eigenvalue on the unit circle must be 1. Next, we have  $X = \text{tr}[X]\sigma$  and hence  $X$  is proportional to  $\sigma$ . Thus, the geometric multiplicity of 1 must be 1: there is a one-dimensional eigenspace. Lastly, we may choose  $\sigma$  as the eigenvector to 1, which is strictly positive definite.

On the other hand, assume  $\Phi$  satisfies the spectral conditions of the proposition, and call  $\tilde{\sigma} \in \mathcal{D}(A)$  the normalized positive definite eigenvector associated to the eigenvalue 1. Then the Jordan decomposition (10.5) of  $\Phi$  gives

$$\Phi = P_1 + \sum_i \mu_i Q_i + D_i$$

where  $P_1$  is the eigenprojection onto  $\tilde{\sigma}$ , and each  $|\mu_i| < 1$ . Then, using the orthogonality relations  $P_1(Q_i + D_i) = (Q_i + D_i)(Q_j + D_j) = 0$  for  $i \neq j$ , we have

$$\Phi^n = P_1 + \sum_i \mu_i^n (Q_i + D_i)^n.$$

The limit  $n \rightarrow \infty$  yields  $\Phi^n \rightarrow P_1$ . Since  $P_1$  is rank-1, as the projection onto a one-dimensional subspace spanned by  $\tilde{\sigma}$ , we have  $P_1(X) = \ell(X)\tilde{\sigma}$  for some linear functional  $\ell$ . Since  $\Phi$  is trace-preserving, we must have  $\ell = \text{tr}$ . Thus, we obtain (10.14) with  $\sigma = \tilde{\sigma}$  and in particular (10.13).  $\square$

We are also interested in a generalization of primitivity called *irreducibility*.

**Definition 10.3.4** (Irreducible map). Let  $\Phi \in \mathcal{B}(\mathcal{B}(A))$  be positive. Then we say  $\Phi$  is *irreducible* if any orthogonal projection  $P \in \mathcal{B}(A)$  satisfies

$$\Phi(P\mathcal{B}(A)P) \subseteq P\mathcal{B}(A)P \implies P \in \{0, \mathbf{1}_A\}. \quad (10.15)$$

*Remark.* The condition given in (10.15) is equivalent to

$$\exists c > 0 \text{ s.t. } \Phi(P) \leq cP \implies P \in \{0, \mathbf{1}_A\}. \quad (10.16)$$

To see this equivalence, note that if  $\Phi(P) \leq cP$ , then for any  $X \in \mathcal{B}(A)$ ,

$$PXP \leq \|X\|_\infty P$$

so  $\Phi(PXP) \leq \|X\|_\infty \Phi(P) \leq c\|X\|_\infty P \in P\mathcal{B}(A)P$ , so  $P\mathcal{B}(A)P$  is invariant under  $\Phi$ . On the other hand, if  $P\mathcal{B}(A)P$  is invariant under  $\Phi$ , then  $\Phi(P) \in P\mathcal{B}(A)P$ , so  $\Phi(P) \leq \|\Phi(P)\|_\infty P$ .

The following characterization in the case of TP irreducible maps parallels the definition of primitivity in Definition 10.3.2. The following can be found as Corollary 6.3 of [Wol12].

**Proposition 10.3.5** (Ergodic sum characterization of irreducibility). *Let  $\Phi \in \mathcal{B}(\mathcal{B}(A))$  be positive and trace-preserving. Then  $\Phi$  is irreducible if and only if there is a full-rank state  $\sigma \in \mathcal{D}(A)$  such that*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \Phi^n(\rho) = \sigma, \quad \forall \rho \in \mathcal{D}(A).$$



Note that Proposition 10.3.5 implies that if  $\Phi \in \mathcal{B}(\mathcal{B}(A))$  is positive, trace-preserving, and primitive, then it is irreducible.

The following spectral characterization can be found in [Wol12, Theorem 6.4], and provides an analog to Proposition 10.3.3 in the case of irreducibility.

**Proposition 10.3.6** (Spectral characterization of irreducibility). *Let  $\Phi \in \mathcal{B}(\mathcal{B}(A))$  be a positive map with spectral radius  $r$ . Then  $\Phi$  is irreducible if and only if  $r$  is a non-degenerate eigenvalue with strictly positive-definite left- and right- eigenvectors, i.e. there exists  $X, Y \in \mathcal{B}(A)$  such that  $\Phi(X) = rX$  and  $\Phi^*(Y) = rY$ .*

*Remark.* Note that if  $\Phi$  is a quantum channel, then  $\text{spr}(\Phi) = 1$  is indeed an eigenvalue with corresponding left-eigenvector  $Y = \mathbf{1}_A > 0$  (since  $\Phi$  is TP), and there always exists a corresponding right-eigenvector  $\sigma \in \mathcal{D}(A)$  due to Proposition 10.3.1. However, while  $\sigma \geq 0$ , in general the state may not be full-rank, and the eigenvalue 1 could be degenerate. An irreducible quantum channel, therefore, is one in which 1 is indeed non-degenerate, and the associated eigenvector  $\sigma$  satisfies  $\sigma > 0$ .

The following property of irreducible maps will prove useful later on.

**Proposition 10.3.7** (Irreducible maps preserve faithfulness). *If  $\Phi \in \mathcal{B}(\mathcal{B}(A))$  is positive and irreducible, then  $\Phi(\mathcal{B}_{++}(A)) \subseteq \mathcal{B}_{++}(A)$ .*

*Proof.* Assume for the sake of contradiction that for some  $X \in \mathcal{B}_{++}(A)$  we have that  $\Phi(X)$  is not full-rank. Let  $\pi \in \mathcal{B}(A)$  be the projection onto the support of  $\Phi(X)$ . Then  $\pi \leq cX$  for some constant  $c > 0$ . Hence, by positivity,  $\Phi(\pi) \leq c\Phi(X) \leq c'\pi$  for some  $c' > 0$ . Since  $\pi \notin \{0, \mathbf{1}_A\}$ , this contradicts the irreducibility of  $\Phi$  via (10.16).  $\square$

The study of the peripheral spectral properties of irreducible maps is the subject of the non-commutative Perron-Frobenius theory for irreducible completely positive maps; see [EH78], [FP09], or [Wol12, Section 6.2]. See also [HJPR18, Appendix A] for a summary of this theory and extensions to deformations of irreducible CPTP maps. Together with the Jordan decomposition (10.5), this theory provides a useful decomposition of irreducible quantum channels. In the next proposition, we recall this decomposition and provide a minimal set of quantities needed to construct such a map. This will prove useful in Chapter 11 for constructing examples of irreducible quantum channels which are not eventually entanglement-breaking.

**Proposition 10.3.8** (Irreducibility via a decomposition). *Consider*

1. An integer  $z \in \{1, \dots, d_{\mathcal{H}}\}$ ,

2. An orthogonal resolution of the identity  $\{p_n\}_{n=0}^{z-1}$ , i.e.,  $\sum_{n=0}^{z-1} p_n = \mathbf{1}$  and  $p_n^\dagger = p_n^2 = p_n$  for each  $n$ ,
3. A faithful state  $\sigma$  such that  $[\sigma, p_n] = 0$  and  $\text{tr}[\sigma p_n] = \frac{1}{z}$ , for each  $n = 0, \dots, z-1$ ,
4. A linear map  $\Phi_Q$  such that:

- (a)  $\text{spr}(\Phi_Q) < 1$
- (b)  $J(\Phi_Q) \geq -z(\sigma \otimes \mathbf{1}) L_1$ , where for  $k = 0, \dots, z-1$ , we define  $L_k := \sum_{n=0}^{z-1} p_{n-k} \otimes p_n$  where the subscripts are taken modulo  $z$ .
- (c) We have

$$\Phi_Q(\sigma p_n) = \Phi_Q^*(p_n) = 0, \quad \forall n = 0, 1, \dots, z-1. \quad (10.17)$$

Let

$$\Phi := \sum_{n=0}^{z-1} \theta^n P_n + \Phi_Q \quad (10.18)$$

where  $P_n(\cdot) = \text{tr}[u^{-n} \cdot] u^n \sigma$  for  $u := \sum_{k=0}^{z-1} \theta^k p_k$  and  $\theta := \exp(2i\pi/z)$ . Then  $\Phi$  is an irreducible quantum channel. On the other hand, any irreducible quantum channel  $\Phi$  can be decomposed as (10.18) for some choices of  $z$ ,  $\{p_n\}_{n=0}^{z-1}$ ,  $\sigma$ , and  $\Phi_Q$  as in (1)–(4). Moreover, in either case,  $\sigma$  is the unique fixed point<sup>1</sup> of  $\Phi$ ;  $P_n(\cdot)$  are its peripheral eigenprojections, associated to eigenvalues  $\theta^n$  and eigenvectors  $u^n \sigma$ ; and, for any  $j, k = 0, \dots, z-1$ , we have the intertwining relations

$$\Phi(p_j X p_k) = p_{j-1} \Phi(X) p_{k-1}, \quad \text{and} \quad \Phi^*(p_j X p_k) = p_{j+1} \Phi^*(X) p_{k+1}, \quad \forall X \in \mathcal{B}(\mathcal{H}) \quad (10.19)$$

where the subscripts are interpreted modulo  $z$ . Additionally, for  $\Phi_P := \sum_{n=0}^{z-1} \theta^n P_n$ ,

$$J(\Phi_P^k) = \hat{J}_k := z(\sigma \otimes \mathbf{1}) L_k = \sum_{m=0}^{z-1} \text{tr}[p_m] \frac{p_{m-k} \sigma p_{m-k}}{\text{tr}[p_{m-k} \sigma]} \otimes \frac{p_m}{\text{tr}[p_m]}. \quad (10.20)$$

*Proof.* Let us note that (4c) is equivalent to the property that

$$\Phi_Q \circ P_j = P_j \circ \Phi_Q = 0, \quad \forall j = 0, \dots, z-1. \quad (10.21)$$

To see this, note that the generalized discrete Fourier transform

$$\mathcal{F} : \bigoplus_{j=0}^{z-1} \mathcal{B}(\mathcal{H}) \rightarrow \bigoplus_{j=0}^{z-1} \mathcal{B}(\mathcal{H})$$

---

<sup>1</sup>up to a multiplicative constant

given by  $\mathcal{F}((X_0, \dots, X_{z-1})) = (Y_0, \dots, Y_{z-1})$  for  $Y_n = \sum_{j=0}^{z-1} \theta^{nj} X_j$  is an invertible linear transformation, with inverse  $\mathcal{F}^{-1}((Y_n)_{n=0}^{z-1}) = \frac{1}{z} \mathcal{F}((Y_{z-\ell})_{\ell=0}^{z-1})$ . All indices are taken mod  $z$ . Next, using the definition of the  $P_n$ , (10.21) is equivalent to

$$\Phi_Q(u^n \sigma) = 0, \quad \Phi_Q^*(u^n) = 0, \quad \forall j = 0, \dots, z-1.$$

Since  $\vec{0} = (\Phi_Q(u^n \sigma))_{n=0}^{z-1} = \mathcal{F}((\Phi_Q(p_j \sigma))_{j=0}^{z-1})$  and  $\vec{0} = (\Phi_Q(u^n))_{n=0}^{z-1} = \mathcal{F}((\Phi_Q(p_j))_{j=0}^{z-1})$ , and  $\mathcal{F}$  has trivial kernel, (10.21) implies (4c). The converse follows similarly. Equation (10.19) follows from a simple computation. The fact that an irreducible map can be decomposed as (10.18) for some choices of  $z$ ,  $\{p_n\}_{n=0}^{z-1}$ ,  $\sigma$  and  $\Phi_Q$  as in (1)-(4b) and (10.21) is not new, and we refer to [Wol12, Section 6.2] for more details. We believe however that the forward implication is novel, however, and include the proof below.

Let us show that given  $z$ ,  $\{p_n\}_{n=0}^{z-1}$ ,  $\sigma$ , and  $\Phi_Q$ , the decomposition (10.18) gives an irreducible quantum channel. Note, by the definition of  $P_n$ , for any  $X \in \mathcal{B}(\mathcal{H})$ ,

$$P_j \circ P_k(X) = \text{tr}[u^{-k} X] \text{tr}[u^{k-j} \sigma] u^j \sigma = \delta_{jk} P_j(X)$$

using  $\text{tr}[\sigma p_n] = \frac{1}{z}$  and the formula

$$\sum_{n=0}^{z-1} \theta^{nm} = \begin{cases} z & m = zk \text{ for some } k \in \mathbb{Z} \\ 0 & \text{otherwise.} \end{cases} \quad (10.22)$$

Since  $P_0(X) = \text{tr}[X] \sigma$ , we have for  $j \neq 0$ ,  $0 = P_0 \circ P_j(X) = \text{tr}[P_j(X)] \sigma$ , yielding that  $P_j$  is trace-annihilating:  $\text{tr}[P_j(X)] = 0$  for all  $X \in \mathcal{B}(\mathcal{H})$ . In the same way, using assumption 4c,  $\Phi_Q$  is trace-annihilating. Thus,  $\Phi = P_0 + \sum_{n=1}^{z-1} \theta^n P_n + \Phi_Q$  is trace-preserving.

Next, we prove (10.20), which will prove  $\Phi$  is CP via assumption 4b. For  $\Phi_P := \sum_{m=0}^{z-1} \theta^m P_m$ , we have  $\Phi_P^k = \sum_{m=0}^{z-1} \theta^{km} P_m$ . Then, for any  $X \in \mathcal{B}(\mathcal{H})$ , we have the discrete Fourier-type computation,

$$\begin{aligned} \Phi_P^k(X) &= \sum_{m=0}^{z-1} \theta^{km} \text{tr}[u^{-m} X] u^m \sigma = \sum_{m,n,\ell=0}^{z-1} \theta^{km} \text{tr}[\theta^{-mn} p_n X] \theta^{\ell m} p_\ell \sigma \\ &= \sum_{m,n,\ell=0}^{z-1} \theta^{m(k-n+\ell)} \text{tr}[p_n X] p_\ell \sigma = \sum_{n,\ell=0}^{z-1} z \delta_{\ell=n-k} \text{tr}[p_n X] p_\ell \sigma \\ &= z \sum_{n=0}^{z-1} \text{tr}[p_n X] p_{n-k} \sigma \end{aligned} \quad (10.23)$$

using (10.22). Next, let  $\{|i\rangle\}_{i=0}^{d_{\mathcal{H}}-1}$  be an orthonormal basis of  $\mathcal{H}$  such that the first  $\text{rank}(p_0)$  elements are a basis for  $p_0\mathcal{H}$ , the next  $\text{rank}(p_1)$  elements are a basis for  $p_2\mathcal{H}$ , and so on. We have  $p_0 = \sum_{i=0}^{\text{rank}(p_0)-1} |i\rangle \langle i|$ , and so forth. Thus,

$$\begin{aligned} J(\Phi_P^k) &= \sum_{i,j=0}^{d-1} P^k(|i\rangle \langle j|) \otimes |i\rangle \langle j| = z \sum_{i,j}^{z-1} \sum_{n=0}^{z-1} \text{tr}[p_n |i\rangle \langle j|] p_{n-k} \sigma \otimes |i\rangle \langle j| \\ &= z \sum_{n=0}^{z-1} \sum_{i=0}^{d-1} \langle i| p_n |i\rangle p_{n-k} \sigma \otimes |i\rangle \langle i| = z \sum_{n=0}^{z-1} \sigma p_{n-k} \otimes p_n \\ &= z(\sigma \otimes \mathbf{1}) L_k. \end{aligned} \quad (10.24)$$

In particular,  $J(\Phi_P) = z(\sigma \otimes \mathbf{1}) L_1$ . Thus, by assumption 4b,

$$J(\Phi) = J(\Phi_P) + J(\Phi_Q) \geq J(\Phi_P) - z(\sigma \otimes \mathbf{1}) L_1 = 0$$

and hence  $\Phi$  is CP. Since  $\Phi$  is CPTP, we can use Proposition 10.3.5 to prove  $\Phi$  is irreducible. We have

$$\frac{1}{M} \sum_{n=0}^{M-1} \Phi^n = P_0 + \frac{1}{M} \sum_{m=1}^{z-1} \frac{1 - \theta^{Mm}}{1 - \theta^m} P_m + \frac{1}{M} \sum_{n=0}^{M-1} \Phi_Q^n$$

using the geometric series  $\sum_{n=0}^{M-1} \theta^{nm} = \frac{1 - \theta^{mM}}{1 - \theta^m}$  for  $m \neq 0$ , which is valid as  $\theta^m \neq 1$ . Since  $P_0[X] = \text{tr}[X]\sigma$ , it remains to show that the latter two terms vanish in the limit  $M \rightarrow \infty$ . In fact, since  $\sum_{m=1}^{z-1} \frac{1 - \theta^{Mm}}{1 - \theta^m} P_m$  is bounded in norm uniformly in  $M$ , the second term vanishes asymptotically. Next, since  $\ell := \text{spr}(\Phi_Q) < 1$  by assumption 4a, for  $\varepsilon = \frac{1-\ell}{2} > 0$ , Gelfand's theorem gives that there is  $n_0 > 0$  such that (in any matrix norm  $\|\cdot\|$ ), for all  $n \geq n_0$ ,

$$\|\Phi_Q^n\| \leq (\ell + \varepsilon)^n < 1.$$

We may write

$$\frac{1}{M} \sum_{n=0}^{M-1} \Phi_Q^n = \frac{1}{M} \sum_{n=0}^{n_0} \Phi_Q^n + \frac{1}{M} \sum_{n=n_0+1}^{M-1} \Phi_Q^n.$$

Since  $\sum_{n=0}^{n_0} \Phi_Q^n$  is bounded in norm independently of  $M$ , the first term vanishes asymptotically; the second term is bounded in norm by the triangle inequality and the geometric series  $\sum_{n=0}^{\infty} (\ell + \varepsilon)^n = \frac{1}{1 - (\ell + \varepsilon)}$ . Thus, the limit

$$\lim_{M \rightarrow \infty} \frac{1}{M} \sum_{n=0}^{M-1} \Phi^n = P_0$$

holds in any norm. In particular, by Proposition 10.3.5  $\Phi$  is irreducible.  $\square$

*Remarks.*

- The proof of Proposition 10.3.5 given in [Wol12, Corollary 6.3] relies on the reverse direction of Proposition 10.3.8, while the proof of the forward direction of Proposition 10.3.8 relies on Proposition 10.3.5. Thus, circular logic is avoided by proving first the reverse direction of Proposition 10.3.8, then Proposition 10.3.5, then the forward direction of Proposition 10.3.8.
- Under the assumptions of the proposition,  $z = 1$  if and only if  $\Phi$  is primitive. This follows from the spectral characterization of primitivity given in Proposition 10.3.3.
- The intertwining property (10.19) holds for  $\Phi$  and  $\Phi_P$  (which itself is an irreducible channel), and therefore for  $\Phi_Q$ . This implies  $J(\Phi_Q) = L_1 J(\Phi_Q) L_1$ , i.e., the Choi matrix of  $\Phi_Q$  is supported on the same subspace as that of  $\Phi_P$ . Additionally, [FP09, Theorem 5.4] shows that for any Kraus decomposition  $\{V_i\}_i \subseteq \mathcal{B}(A)$ , i.e. such that  $\Phi(X) = \sum_i V_i X V_i^*$  holds for all  $X \in \mathcal{B}(A)$ , the intertwining relations

$$p_m V_i = V_i p_{m+1} \quad \forall m, i \quad (10.25)$$

also hold.

- Given a map  $\Phi_Q$  which intertwines with  $\{p_n\}_{n=0}^{z-1}$ , a sufficient condition for  $\Phi_Q \geq -z(\sigma \otimes \mathbf{1}) L_1 \equiv -J(\Phi_P)$  is given by

$$\|\Phi_Q\|_2 = \|J(\Phi_Q)\|_2 \leq z \lambda_{\min}(\sigma), \quad (10.26)$$

since in that case

$$\begin{aligned} J(\Phi_Q) &\geq -\text{spr}(\Phi_Q) L_1 \geq -\|J(\Phi_Q)\|_2 L_1 \\ &\geq -z \lambda_{\min}(\sigma) L_1 \geq -z L_1 (\sigma \otimes \text{id}) L_1 = -J(\Phi_P). \end{aligned} \quad (10.27)$$

- The matrix  $\hat{J}_k$  is separable, and thus  $\Phi_P^k$  is entanglement-breaking, for any  $k \geq 1$ .
- If  $z > 1$ , then  $\frac{1}{d_{\mathcal{H}}} \hat{J}_k$  does not have full support. Thus,  $\frac{1}{d_{\mathcal{H}}} \hat{J}_k$  is on the boundary of the set of density matrices, and thus on the boundary of SEP and PPT as well. In fact, we can say something stronger than this: whenever  $z > 1$ , there exist entangled density matrices arbitrarily close to each  $\frac{1}{d_{\mathcal{H}}} \hat{J}_k$ ,  $k = 0, 1, \dots, z-1$ . To see this, note that  $\frac{1}{d_{\mathcal{H}}} \hat{J}_k \in L_k \mathcal{D}(\mathcal{H} \otimes \mathcal{H}) L_k$ . However, we can construct entangled

states in  $L_j \mathcal{D}(\mathcal{H} \otimes \mathcal{H}) L_j$  for any  $j \neq k$ . For instance, let  $|0\rangle \in p_0 \mathcal{H}$ ,  $|1\rangle \in p_1 \mathcal{H}$ ,  $|j_0\rangle \in p_{-j} \mathcal{H}$ , and  $|j_1\rangle \in p_{1-j} \mathcal{H}$  be normalized vectors. Then

$$|\Omega_j\rangle := \frac{1}{\sqrt{2}}(|j_0\rangle \otimes |0\rangle + |j_1\rangle \otimes |1\rangle)$$

is (local-unitarily equivalent to) a Bell state, and has

$$L_j |\Omega_j\rangle = \sum_{n=0}^{z-1} (p_{n-j} \otimes p_n) |\Omega_j\rangle = |\Omega_j\rangle,$$

and thus  $L_j |\Omega_j\rangle \langle \Omega_j| L_j \in L_j \mathcal{D}(\mathcal{H} \otimes \mathcal{H}) L_j$ . Thus, for any  $t \geq 0$ ,

$$(1-t) \frac{1}{d_{\mathcal{H}}} \hat{J}_k + t |\Omega_j\rangle \langle \Omega_j| = (1-t) \frac{1}{d_{\mathcal{H}}} \hat{J}_k \Big|_{L_k} \oplus t |\Omega_j\rangle \langle \Omega_j| \Big|_{L_j}$$

is an entangled density matrix, and can be made arbitrarily close to  $\frac{1}{d_{\mathcal{H}}} \hat{J}_k$  by sending  $t \rightarrow 0$ .

One can relate the decompositions given in Proposition 10.3.8 and (10.12), as follows. The following result is straightforward, but apparently novel (first appearing in [HRS20]).

**Proposition 10.3.9.** *The following are equivalent:*

1. *There exists a decomposition (10.11) such that  $\Phi$  satisfies (10.12) with  $\mathcal{K}_0 = \{0\}$  and  $d_i = 1$  for each  $i$ , and  $\pi$  is a  $K$ -cycle*
2.  *$\Phi$  is irreducible.*

*Proof.* First, assume  $\Phi$  is irreducible, and adopt the notation of Proposition 10.3.8. Then

$$\tilde{N}(\Phi) = \text{span}\{u^n \sigma : n = 0, \dots, z-1\}.$$

Let  $X \in \tilde{N}(\Phi)$  be given by  $X = \sum_{j=0}^{z-1} \lambda_j u^j \sigma$  for some  $\lambda_j \in \mathbb{C}$ . Since  $\sum_{n=0}^{z-1} p_n = \mathbf{1}$ , we have

$$X = \sum_{n,j=0}^{z-1} \lambda_j p_n u^j \sigma = \sum_{n,j=0}^{z-1} \lambda_j \theta^{nj} p_n \sigma = \bigoplus_{n=0}^{z-1} \left( \sum_{j=0}^{z-1} \lambda_j \theta^{nj} \right) \sigma|_{p_n} \quad (10.28)$$

where  $\sigma|_{p_n}$  is  $\sigma$  restricted to the subspace  $p_n \mathcal{H}$ , and the direct sum decomposition is with respect to the decomposition  $\mathcal{H} = \bigoplus_{n=0}^{z-1} p_n \mathcal{H}$ . Note moreover,  $\sum_{j=0}^{z-1} \lambda_j \theta^{nj}$  is the  $n$ th coefficient of the discrete Fourier transform  $F_z : \mathbb{C}^z \rightarrow \mathbb{C}^z$  of the vector  $\lambda := (\lambda_j)_{j=0}^{z-1}$ .

Since the Fourier transform is invertible, as  $\lambda$  ranges over  $\mathbb{C}^z$ , the vector of Fourier coefficients range over  $\mathbb{C}^z$  as well. We therefore find

$$\tilde{\mathcal{N}}(\Phi) = \bigoplus_{n=0}^{z-1} \mathbb{C} \sigma|_{p_n}$$

which is a decomposition of the form given by (10.11) with  $\mathcal{K}_0 = \{0\}$  and  $d_n = 1$  for each  $n$ . Moreover, for  $X \in \tilde{\mathcal{N}}(\Phi)$  decomposed as  $X = \bigoplus_{n=0}^{z-1} \gamma_n \sigma|_{p_n} = \sum_{n=0}^{z-1} \gamma_n p_n \sigma p_n$ , we have

$$\Phi(X) = \sum_{n=0}^{z-1} \gamma_n \Phi(p_n \sigma p_n) = \sum_{n=0}^{z-1} \gamma_n p_{n-1} \Phi(\sigma) p_{n-1} = \sum_{n=0}^{z-1} \gamma_n p_{n-1} \sigma p_{n-1}$$

by (10.19). Thus, (10.12) holds, where  $\pi$  is the cyclic permutation  $k \rightarrow k+1$ .

On the other hand, assume we are given such a decomposition,

$$\tilde{\mathcal{N}}(\Phi) = \bigoplus_{i=0}^{K-1} \mathbb{C} \tau_i$$

for some states  $\tau_i \in \mathcal{D}_+(\mathcal{K}_i)$  where  $\mathcal{H} = \bigoplus_{i=0}^{K-1} \mathcal{K}_i$ , such that<sup>2</sup>

$$\Phi(X) = \bigoplus_{i=0}^{K-1} \gamma_{i+1} \tau_i, \quad \text{for} \quad X = \bigoplus_{i=0}^{K-1} \gamma_i \tau_i.$$

Define  $\sigma := \bigoplus_{i=0}^{K-1} \frac{1}{z} \tau_i$ . Then by (10.12),

$$\Phi(\sigma) = \bigoplus_{i=0}^{K-1} \frac{1}{z} \tau_i = \sigma$$

as suggested by the name. On the other hand, assume  $X \in \mathcal{B}(\mathcal{H})$  is also invariant under  $\Phi$ . Then  $X \in \tilde{\mathcal{N}}(\Phi)$ , so  $X = \bigoplus_{i=0}^{K-1} \lambda_i \tau_i$  for some  $\lambda_i \in \mathbb{C}$ . But then

$$\Phi(X) = \bigoplus_{i=0}^{K-1} \lambda_{i+1} \tau_i = X \implies \lambda_i = \lambda_{i+1} \forall i.$$

Thus,  $X$  is proportional to  $\sigma$ , and  $\Phi$  must have a unique invariant state. Since  $\sigma$  is additionally full-rank (by construction), we conclude that  $\Phi$  is irreducible by Proposition 10.3.6.  $\square$

<sup>2</sup>If  $\pi$  is a cycle, we may reorder the index of the direct sum so that  $\pi$  maps  $i$  to  $i+1$ .





# Chapter 11

## Characterizing EEB channels

### 11.1 Introduction through (counter)examples

In this chapter, we are interested in characterizing which quantum channels are eventually entanglement-breaking (EEB). Here, we will explore this question via a sequence of motivating examples and counterexamples. Recall that a linear map  $\Phi$  is called *eventually entanglement-breaking* (EEB) if  $\Phi^n$  is entanglement-breaking for some  $n \in \mathbb{N}$ .

Let's first see an example of an EEB map: a primitive quantum channel. As defined in Definition 10.3.2, a quantum channel  $\Phi$  on  $\mathcal{B}(\mathcal{H})$  to be *primitive* if there exists a full-rank state  $\sigma \in \mathcal{D}(\mathcal{H})$  such that we have the asymptotic convergence

$$\Phi^n(\rho) \xrightarrow{n \rightarrow \infty} \sigma, \quad \forall \rho \in \mathcal{D}(\mathcal{H}). \quad (11.1)$$

Let us see that if  $\Phi \in \mathcal{B}(\mathcal{B}(A) \rightarrow \mathcal{B}(A))$  is primitive, it is indeed EEB. Let  $R \cong A$  be a reference system, and set  $d := d_A = d_R$ . Then to check if  $\Phi^n$  is EB, we compute its Choi state,

$$\begin{aligned} \Phi^n \otimes \text{id}_R(\Omega_{A:R}) &= \frac{1}{d} \sum_{ij} \Phi^n(|i\rangle \langle j|_A) \otimes |i\rangle \langle j|_R \\ &\rightarrow \frac{1}{d} \sum_{ij} \text{tr}[|i\rangle \langle j|_A] \sigma_A \otimes |i\rangle \langle j|_R \\ &= \sigma_A \otimes \frac{1}{d} \sum_{ij} \delta_{ij} |i\rangle \langle j|_R \\ &= \sigma_A \otimes \frac{1}{d} \sum_i |i\rangle \langle i|_R = \sigma_A \otimes \tau_R \end{aligned}$$

where  $\tau_R := \frac{1}{d_R} \mathbf{1}_R$ . We see this is indeed a product state. However, this does not yet complete the proof, since we need to show that  $\Phi^n$  is entanglement-breaking for finite  $n$ , not only in the limit  $n \rightarrow \infty$ . To complete the proof, we use the following corollary to [GB02, Theorem 1].

**Lemma 11.1.1.** *Let  $\omega_A, \sigma_B > 0$ . Then  $\omega_A \otimes \sigma_B + \Delta_{AB}$  is separable for any Hermitian operator  $\Delta_{AB}$  such that  $\|\Delta_{AB}\|_2 \leq \lambda_{\min}(\sigma_A) \lambda_{\min}(\omega_B)$ , where  $\lambda_{\min}(\omega_A)$ , resp.  $\lambda_{\min}(\sigma_B)$ , stands for the smallest eigenvalue of  $\omega_A$ , resp.  $\sigma_B$ .*

*Proof.* Theorem 1 of [GB02] shows that if  $\|\tilde{\Delta}_{AB}\|_2 \leq 1$  then  $\mathbf{1}_A \otimes \mathbf{1}_B + \tilde{\Delta}_{AB} \in \text{SEP}(A : B)$ . Since

$$\omega_A \otimes \sigma_B + \Delta_{AB} \in \text{SEP} \iff \mathbf{1}_A \otimes \mathbf{1}_B + \omega_A^{-1/2} \otimes \sigma_B^{-1/2} \Delta_{AB} \omega_A^{-1/2} \otimes \sigma_B^{-1/2} \in \text{SEP}(A : B)$$

the proof is concluded by the fact that

$$\|\omega_A^{-1/2} \otimes \sigma_B^{-1/2} \Delta_{AB} \omega_A^{-1/2} \otimes \sigma_B^{-1/2}\|_2 \leq \|\omega_A^{-1}\|_\infty \|\sigma_B^{-1}\|_\infty \|\Delta_{AB}\|_2.$$

□

*Remark.* This short and simple proof is due to an anonymous referee of the paper [HRS20].

Returning to our primitive map  $\Phi$ , we saw that

$$\Phi^n \otimes \text{id}_R(\Omega_{A:R}) \xrightarrow{n \rightarrow \infty} \sigma_A \otimes \tau_R.$$

Hence, for  $n$  large enough, it holds that

$$\|\Phi^n(\Omega_{A:R}) - \sigma_A \otimes \tau_R\|_2 \leq \lambda_{\min}(\sigma_A) \lambda_{\min}(\tau_R) = \frac{1}{d} \lambda_{\min}(\sigma_A).$$

Thus, Lemma 11.1.1 proves that  $\Phi^n(\Omega_{A:R}) \in \text{SEP}(A : R)$  for  $n$  sufficiently large; hence,  $\Phi$  is EEB.

Now that we have found an example of a class of EEB quantum channels, let us speculate this is the only such class<sup>1</sup>.

*Wrong conjecture 1:*  $\Phi$  is EEB if and only if it is primitive.

*Counterexample:* Consider the simple quantum channel

$$\Phi(\rho) = \text{tr}(\rho) |0\rangle \langle 0|.$$

---

<sup>1</sup>This turns out to be correct in the simpler case of continuous-time Markov quantum semigroups [HRS20]

Then  $\Phi$  has a rank-1 invariant state (and thus is not primitive), but is also entanglement-breaking and hence eventually entanglement-breaking.

What if we restrict our considerations to channels with a full-rank invariant state?

*Wrong conjecture 2:* If  $\Phi$  has a full-rank invariant state, then it is EEB if and only if it is primitive.

*Counterexample:* Let  $\mathcal{H} = \mathcal{H}^1 \oplus \mathcal{H}^2$ , and  $\Phi = \Phi_1 \oplus \Phi_2$ . That is, for a state  $\rho \in \mathcal{D}(\mathcal{H})$ ,

$$\Phi \left( \begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{pmatrix} \right) := \begin{pmatrix} \Phi_1(\rho_{11}) & 0 \\ 0 & \Phi_2(\rho_{22}) \end{pmatrix}$$

where each  $\rho_{ij}$  is a block-matrix, namely the part of  $\rho$  mapping  $\mathcal{H}^j \rightarrow \mathcal{H}^i$ . If  $\Phi_1$  satisfies (11.1) with some  $\sigma_1$  and  $\Phi_2$  satisfies (11.1) with some state  $\sigma_2$ , then similarly as before,  $\Phi$  is EEB. However,

$$\Phi^n(\rho) \xrightarrow{n \rightarrow \infty} \text{tr}[\rho_{11}]\sigma_1 \oplus \text{tr}[\rho_{22}]\sigma_2,$$

so  $\Phi$  does not satisfy (11.1): the final state still depends on the input.

In this example, however,  $\Phi$  is the direct sum of primitive channels. Let us allow that too.

*Wrong conjecture 3:* If  $\Phi$  has a full-rank invariant state, then it EEB if and only if it is primitive or the direct sum of primitive channels.

*Counterexample:* Consider if for some  $k$ ,  $\Phi^k$  is the sum of primitive channels, but  $\Phi$  itself is not. For example,

$$\Phi \left( \begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{pmatrix} \right) := \begin{pmatrix} \text{tr}(\rho_{22})\sigma_1 & 0 \\ 0 & \text{tr}(\rho_{11})\sigma_2 \end{pmatrix}$$

for some full-rank states  $\sigma_1, \sigma_2$ . Then

$$\Phi^2 \left( \begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{pmatrix} \right) = \begin{pmatrix} \text{tr}(\rho_{11})\sigma_1 & 0 \\ 0 & \text{tr}(\rho_{22})\sigma_2 \end{pmatrix}$$

is the direct sum of primitive channels. Thus,  $\Phi^2$  is eventually entanglement-breaking, as in the previous example, and so  $\Phi$  is EEB too.

It turns out the next logical conjecture, that if  $\Phi$  has a full-rank invariant state, then it is EEB if and only if it is *eventually* primitive or the direct sum of primitive channels, is correct, as shown by Theorem 11.3.8.

## 11.2 Entanglement classes of quantum channels

A map  $\Phi : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H})$  is of *positive partial transpose* (PPT) if  $(\mathcal{T} \circ \Phi) \otimes \text{id}_{\mathcal{H}}$  is a positive operator, where  $\mathcal{T}$  is the partial transpose w.r.t. to some basis. The class of PPT channels on  $\mathcal{H}$  is called  $\text{PPT}(\mathcal{H})$ , and the class of entanglement-breaking channels on  $\mathcal{H}$  is denoted  $\text{EB}(\mathcal{H})$ . As the set of separable bipartite states is a subset of the set of bipartite states with a positive partial transpose, it follows that  $\text{EB}(\mathcal{H}) \subseteq \text{PPT}(\mathcal{H})$ .

Given a quantum channel  $\Phi : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H})$ , the sequence  $\{\Phi^n\}_{n \in \mathbb{N}}$  is called a *discrete time quantum Markov semigroup* (discrete time QMS). Here, semigroup refers simply to the property that  $\Phi^{n+m} = \Phi^n \circ \Phi^m$ . We will often discuss properties of a discrete time QMS  $\{\Phi^n\}_{n \in \mathbb{N}}$  by simply referring to the associated quantum channel  $\Phi$ .

In this chapter, we study the entanglement properties of quantum Markovian evolutions in discrete time. A discrete time QMS  $\{\Phi^n\}_{n \in \mathbb{N}}$  is said to be *eventually entanglement breaking* (EEB) if there exists  $n_0 \in \mathbb{N}$  such that for any  $n \geq n_0$ ,  $\Phi^n$  is entanglement breaking. The class of eventually entanglement breaking Markovian evolutions is denoted by  $\text{EEB}(\mathcal{H})$ . We also say a quantum channel  $\Phi$  is eventually entanglement breaking if the discrete QMS  $\{\Phi^n\}_{n=1}^{\infty}$  is eventually entanglement breaking. On the other hand, Markovian evolutions which are not entanglement breaking at any finite time are called *entanglement saving*, using language introduced by Lami and Giovannetti [LG16]; the class of entanglement saving Markovian evolutions is denoted by  $\text{ES}(\mathcal{H})$ . Thus, the set of all discrete-time Markovian evolutions decomposes into two disjoint classes,

$$\text{EEB}(\mathcal{H}) \sqcup \text{ES}(\mathcal{H}). \quad (11.2)$$

Lami and Giovannetti also introduce the notion of *asymptotically entanglement saving* evolutions in the discrete-time case. They showed that every discrete time QMS has at least one limit point, and either all of the limit points of a discrete time QMS  $\{\Phi^n\}_{n=1}^{\infty}$  are entanglement breaking, or none of them are. They term the latter case as asymptotically entanglement saving, and we denote the set of asymptotically entanglement saving evolutions on  $\mathcal{H}$  as  $\text{AES}(\mathcal{H})$ . In analogy, we call the former case by *asymptotically entanglement breaking*, denoted  $\text{AEB}(\mathcal{H})$ . Thus, the set of discrete time QMS on  $\mathcal{H}$  decomposes into the disjoint classes

$$\text{AES}(\mathcal{H}) \sqcup \text{AEB}(\mathcal{H}). \quad (11.3)$$

It is interesting to compare (11.2) and (11.3). A discrete time QMS  $\{\Phi^n\}_{n=1}^{\infty}$  is AES if the limit points of the sequence  $\{J(\Phi^n)\}_{n=1}^{\infty}$  are all entangled. Since  $J(\Phi^{n+1}) = \Phi \otimes \text{id}(J(\Phi^n))$ ,

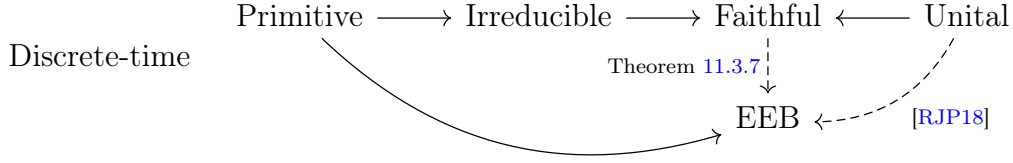


Fig. 11.1 Relations between classes of quantum Markov semigroups, in which arrows represent subsets; e.g., primitive discrete-time quantum Markov semigroups are a subset of irreducible discrete-time quantum Markov semigroups. The dashed arrows indicate relations which only hold for quantum Markov semigroups  $\{\Phi^n\}_{n=1}^\infty$  associated to a PPT channel  $\Phi$ . The arrows without annotations are standard and are discussed in the text.

if  $J(\Phi^{n+1})$  is entangled,  $J(\Phi^n)$  must be as well. In particular, if  $\{\Phi^n\}_{n=1}^\infty \in \text{AES}(\mathcal{H})$ , then  $J(\Phi^n)$  is entangled for every  $n$ , and the discrete time QMS is entanglement saving. So we see  $\text{AES}(\mathcal{H}) \subseteq \text{ES}(\mathcal{H})$ . However, a priori, an entanglement saving QMS could be asymptotically entanglement breaking: at any finite  $n$ ,  $J(\Phi^n)$  could be entangled, but in the limit,  $J(\Phi^n)$  could be in the set of separable states (though necessarily on the boundary). We therefore define  $\text{EB}_\infty(\mathcal{H}) = \text{AEB}(\mathcal{H}) \cap \text{ES}(\mathcal{H})$ , the set of discrete time QMS which are asymptotically entanglement breaking, but not entanglement breaking for any finite  $n$ . With this notion, we may relate (11.2) and (11.3). We have the disjoint decomposition of the set of all discrete time QMS, denoted  $\text{dQMS}(\mathcal{H})$ , satisfies

$$\text{dQMS}(\mathcal{H}) = \overbrace{\text{EEB}(\mathcal{H}) \sqcup \text{EB}_\infty(\mathcal{H})}^{\text{AEB}(\mathcal{H})} \sqcup \underbrace{\text{AES}(\mathcal{H})}_{\text{ES}(\mathcal{H})}. \quad (11.4)$$

*Remark.* The isomorphism between bipartite states and CP maps given by the Choi matrix and the equivalence between EB maps and separable states discussed in Section 10.2.2 allows us to directly translate results on the entanglement loss of CP maps to statements about the separability of a bipartite state. We will mostly state our results in the picture of CPTP maps (see also Section 11.3.3 for a method to remove the TP assumption), but it should be straightforward to derive the corresponding statements for bipartite states.

Some of the definitions introduced in this chapter, along with a preview of some of the results, are depicted diagrammatically in Figure 11.1.

Let us now consider an example to illustrate these definitions and show that all sets arise naturally in physical systems.

**Example 11.2.1.** Consider a discrete time quantum Markov semigroup  $\{\Phi^n\}_{n \in \mathbb{N}}$  associated to a *repeated interaction system*, in which a system  $\mathcal{S}$  interacts with a chain of

identical *probes*  $\mathcal{E}_k$ , one at a time, for a duration  $\tau$ . During the interaction, the dynamics of the system are modeled by a Hamiltonian evolution, and at the times  $(k\tau)_{k \geq 1}$ , the evolution forms a semigroup. In this example, the system and each probe are 2-level systems, with associated Hilbert spaces  $\mathcal{H}_S = \mathcal{H}_E = \mathbb{C}^2$ . We define Hamiltonians  $h_S = E a^* a$  and  $h_E = E_0 b^* b$ , where  $a/a^*$ , resp.  $b/b^*$ , are the annihilation/creation operators for  $\mathcal{S}$ , resp.  $\mathcal{E}$ , and  $E \in \mathbb{R}_{>0}$  (resp.  $E_0 \in \mathbb{R}_{>0}$ ) corresponds to the energy of the excited state of  $\mathcal{S}$  (resp.  $\mathcal{E}$ ). We can express these operators in the (ground state, excited state) basis of each system by

$$a = b = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad a^* = b^* = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad a^* a = b^* b = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

We consider the initial state of each probe to be a *thermal state*,

$$\xi_\beta = \frac{\exp(-\beta h_E)}{\text{tr}[\exp(-\beta h_E)]} = \begin{pmatrix} \frac{1}{1+e^{-\beta E_0}} & 0 \\ 0 & \frac{e^{-\beta E_0}}{1+e^{-\beta E_0}} \end{pmatrix}$$

where  $\beta \in [0, \infty]$  represents the inverse temperature (setting Boltzmann's constant to one). In the case of zero-temperature ( $\beta = \infty$ ), we take

$$\xi_\infty = \lim_{\beta \rightarrow \infty} \xi_\beta = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

We consider an interaction modeling the two systems coupling through their dipoles, in the *rotating wave approximation*. In this setting, the system and each probe interact via the potential  $\lambda v_{\text{RW}} \in \mathcal{B}(\mathcal{H}_S \otimes \mathcal{H}_E)$ , where  $\lambda \geq 0$  is a coupling constant, and

$$v_{\text{RW}} = \frac{u_1}{2}(a^* \otimes b + a \otimes b^*)$$

where  $u_1$  is a constant, which we take to be equal to 1 with units of energy. This is a common approximation in the regime  $|E - E_0| \ll \min\{E, E_0\}$  and  $\lambda \ll |E_0|$ .

The system begins in some state  $\rho^i$ , couples with the first probe (in thermal state  $\xi_\beta$ ), and evolves for a time  $\tau > 0$  according to the unitary operator

$$U := \exp(-i\tau(h_S \otimes \text{id} + \text{id} \otimes h_E + \lambda v_{\text{RW}})).$$

That is,  $\rho^i \otimes \xi_\beta$  evolves to  $U(\rho^i \otimes \xi_\beta)U^*$ . Then we trace out the probe to obtain

$$\rho_1 := \text{tr}_E(U(\rho^i \otimes \xi_\beta)U^*).$$

This process is repeated, and at the end of the  $k$ th step, the system is in the state

$$\rho_k = \text{tr}_{\mathcal{E}}(U(\rho_{k-1} \otimes \xi_{\beta})U^*) = \Phi^k(\rho^i)$$

where  $\Phi$  is the quantum channel given by

$$\Phi(\eta) := \text{tr}_{\mathcal{E}}(U(\eta \otimes \xi_{\beta})U^*) = \Phi^k(\rho^i).$$

What class in the decomposition (11.4) does the discrete time QMS  $\{\Phi^n\}_{n \in \mathbb{N}}$  lie in? To answer that question, we first compute the eigenvectors and eigenvalues of  $\Phi$ , yielding

$$\Phi(\rho_{\beta^*}) = \rho_{\beta^*}, \quad \Phi(a) = \gamma a, \quad \Phi(a^*) = \bar{\gamma} a^*, \quad \Phi(\sigma_z) = |\gamma|^2 \sigma_z$$

where  $\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$  is the Pauli- $z$  matrix,

$$\gamma := e^{-\frac{1}{2}i\tau(E_0+E)} \left( \cos \frac{\tau\nu}{2} + i \frac{(E_0-E)}{\nu} \sin \frac{\tau\nu}{2} \right)$$

for  $\nu := \sqrt{(E_0-E)^2 + \lambda^2}$ , and, defining the rescaled inverse temperature  $\beta^* := \frac{E_0}{E}\beta$ ,

$$\rho_{\beta^*} := \begin{pmatrix} \frac{1}{1+e^{-\beta^*E}} & 0 \\ 0 & \frac{e^{-\beta^*E}}{1+e^{-\beta^*E}} \end{pmatrix} = \xi_{\beta}$$

is a thermal state on  $\mathcal{H}_{\mathcal{S}}$  represented by the same matrix as  $\xi_{\beta}$ . In the case  $\beta = \infty$ , we set  $\rho_{\beta^*} := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \xi_{\infty}$ . Note

$$|\gamma| = \sqrt{1 - \frac{\lambda^2}{\nu^2} \sin^2 \left( \frac{\nu\tau}{2} \right)} \leq 1$$

as required, since  $\Phi$  is a CPTP map. In [LG16, Prop. 23] the authors show that a qubit channel is AES if and only if it is a unitary channel, which in turn is equivalent to it having determinant 1. Thus, we have  $\{\Phi^n\}_{n \in \mathbb{N}} \in \text{AES}(\mathcal{H})$  if and only if  $|\gamma| = 1$ .

To analyze whether or not  $\Phi^n$  is entanglement-breaking, it suffices to check if  $J(\Phi^n)$  is PPT, as  $\Phi$  is a qubit channel. To that end, define the Gibbs factor  $g = \exp(-\beta E_0)$

and partition function  $Z = 1 + g$ . Then we have

$$\rho_{\beta^*} + gZ^{-1}\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \rho_{\beta^*} - Z^{-1}\sigma_z = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Thus, taking  $(|0\rangle, |1\rangle)$  to be the (ground state, excited state) basis for each system,

$$\begin{aligned} (\text{id} \otimes \mathcal{T})J(\Phi^n) &= \sum_{i,j=0}^1 \Phi^n(|i\rangle\langle j|) \otimes |j\rangle\langle i| \\ &= \begin{pmatrix} \Phi^n(\rho_{\beta^*} + Z^{-1}\sigma_z) & \Phi^n(a^*) \\ \Phi^n(a) & \Phi^n(\rho_{\beta^*} - gZ^{-1}\sigma_z) \end{pmatrix} \\ &= \begin{pmatrix} \rho_{\beta^*} + |\gamma|^{2n}Z^{-1}\sigma_z & \bar{\gamma}^n a^* \\ \gamma^n a & \rho_{\beta^*} - |\gamma|^{2n}gZ^{-1}\sigma_z \end{pmatrix} \\ &= \begin{pmatrix} Z^{-1}(g + |\gamma|^{2n}) & 0 & 0 & 0 \\ 0 & Z^{-1}(1 - |\gamma|^{2n}) & \bar{\gamma}^n & 0 \\ 0 & \gamma^n & Z^{-1}g(1 - |\gamma|^{2n}) & 0 \\ 0 & 0 & 0 & Z^{-1}(1 + g|\gamma|^{2n}) \end{pmatrix} \end{aligned}$$

which has eigenvalues

$$Z^{-1}(g + |\gamma|^{2n}), \quad Z^{-1}(1 + g|\gamma|^{2n}), \quad \frac{1 - |\gamma|^{2n}}{2} \pm \frac{1}{2} \sqrt{(1 - |\gamma|^{2n})^2 \left( \frac{g - g^{-1}}{g + g^{-1}} \right)^2 + 4|\gamma|^{2n}}$$

for  $g \in (0, 1]$ , and

$$1, \quad 1, \quad \pm |\gamma|^{2n}$$

for  $g = 0$ . In either case, the eigenvalues only depend on the independent parameters  $|\gamma| \in [0, 1]$  and  $g \in [0, 1]$ . Since  $(\text{id} \otimes \mathcal{T})J(\Phi^n) \geq 0$  is equivalent to  $\Phi^n \in \text{EB}(\mathcal{H})$ , we find

- If  $|\gamma| = 1$ ,  $\{\Phi^n\}_{n \in \mathbb{N}} \in \text{AES}(\mathcal{H})$ . This occurs when  $\nu\tau \in 2\pi\mathbb{Z}$ ; in this case,  $\Phi$  is a unitary channel.
- If  $|\gamma| = 0$ , then  $\Phi \in \text{EB}(\mathcal{H})$ . This occurs when  $E = E_0$ , and  $\nu\tau \in \pi + 2\pi\mathbb{Z}$ . In this case,  $\Phi = \rho_{\beta^*} \text{tr}[\cdot]$ .
- If  $|\gamma| \in (0, 1)$  and  $g = 0$ ,  $\{\Phi^n\}_{n \in \mathbb{N}} \in \text{EB}_\infty(\mathcal{H})$ . In this case,  $\beta = \infty$  and  $\Phi$  has a unique peripheral eigenvalue, namely 1, with non-full-rank invariant state,  $|1\rangle\langle 1|$ .



- If  $|\gamma| \in (0, 1)$  and  $g \in (0, 1]$ , then  $\{\Phi^n\}_{n=1}^\infty \in \text{EEB}(\mathcal{H})$ , and in particular, the minimal  $n$  such that  $\Phi^n \in \text{EB}(\mathcal{H})$  is given by  $n = \max(1, \lceil \frac{1}{2} \frac{\log B(g)}{\log |\gamma|} \rceil)$  where

$$B(g) := \frac{1 + 4g^2 + g^4 - (1 + g^2)\sqrt{1 + 6g^2 + g^4}}{2g^2} \in (0, 3 - 2\sqrt{2}] \quad \text{for } g \in (0, 1].$$

In this case,  $\Phi$  is primitive, with faithful invariant state  $\rho_{\beta^*}$ .

Repeated interaction systems are returned to in Chapter 12 and this example is revisited in Section 12.9.

## 11.3 Which quantum channels are eventually entanglement-breaking?

In order to better characterize discrete-time evolutions for which asking the titular question makes sense, we first need to leave aside those evolutions for which the phenomenon does not occur, that is, evolutions that either destroy entanglement after an infinite amount of time ( $\text{EB}_\infty$ ), or even those of never-vanishing output entanglement (AES).

A large part of this question was already answered in the discrete time case by [LG16]. In that work (Theorem 21), the authors showed that, given a quantum channel  $\Phi : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H})$  with  $\dim(\ker \Phi) < 2(d_{\mathcal{H}} - 1)$ ,  $\{\Phi^n\}_{n \in \mathbb{N}} \in \text{ES}(\mathcal{H})$  if and only if either it has a non-full rank positive fixed point, or the number of peripheral eigenvalues is strictly greater than 1, which itself is equivalent to the existence of  $1 \leq n \leq d_{\mathcal{H}}$  such that  $\Phi^n$  has a non-full rank positive fixed point. In the same paper, the authors showed that if  $\Phi$  has more than  $d_{\mathcal{H}}$  peripheral eigenvalues, then  $\{\Phi^n\}_{n \in \mathbb{N}}$  is asymptotically entanglement saving. These interesting results clearly show the link between the spectral properties of the quantum channel  $\Phi$  and the entanglement properties of the corresponding discrete time quantum Markov semigroup  $\{\Phi^n\}_{n \in \mathbb{N}}$ . In the next subsections, we further develop this intuition. First, we prove the following simple consequence of a result from [LG16].

**Lemma 11.3.1.** *Let  $\Phi$  be a quantum channel on  $\mathcal{H}$  with  $d_{\mathcal{H}}$  peripheral eigenvalues counted with multiplicity and at least one non-zero non-peripheral eigenvalue. Then  $\{\Phi^n\}_{n \in \mathbb{N}} \in \text{ES}(\mathcal{H})$ .*

*Proof.* For any  $N$ ,  $\Phi^N$  has  $d_{\mathcal{H}}$  peripheral eigenvalues and at least one non-peripheral eigenvalue non-zero. Thus, if  $\{\lambda_k\}_{k=1}^{d_{\mathcal{H}}^2}$  are the eigenvalues of  $\Phi$  counted with multiplicity,

we have

$$\|\Phi^N\|_1 \geq \sum_{k=1}^{d^2} |\lambda_k^N| = d_{\mathcal{H}} + \sum_{\lambda_k: |\lambda_k| < 1} |\lambda_k^N| > d_{\mathcal{H}}.$$

The result follows from the fact that a quantum channel  $\Psi : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H})$ ,

$$\|\Psi\|_1 > d_{\mathcal{H}} \implies \Psi \notin \text{EB}(\mathcal{H}) \quad (11.5)$$

which is due to the reshuffling criterion [CW03] (see, e.g. [LG16, eq. (47)]).  $\square$

### 11.3.1 Irreducible evolutions

The previous result suggest looking at channels with less than  $d_{\mathcal{H}}$  peripheral eigenvalues counted with multiplicity in order to characterize EEB. Since primitive channels are eventually entanglement-breaking, a natural next step is to consider the wider class of irreducible channels introduced in Section 10.3.

As a first step, Proposition 10.3.8 shows that the peripheral eigenvectors of irreducible channels  $\Phi$  commute. In Theorem 32 of [LG16], the authors show that asymptotically entanglement saving channels are characterized by the fact that they possess at least two noncommuting phase points. This implies  $\{\Phi^n\}_{n \in \mathbb{N}} \in \text{AEB}(\mathcal{H})$ , which generalizes Corollary 6.1 of [RJP18] to the non-unital case. There, the authors show that a unital irreducible quantum channel is AEB if and only if its phase space is commutative.

Next, the intertwining property (10.19) of irreducible maps is very useful for understanding their entanglement breaking properties. In fact, a slight generalization of this property will prove useful. Recall that an orthogonal resolution of the identity is a collection of orthogonal projections which sum to the identity.

**Definition 11.3.2** ( $(\{p_i\}_{i=0}^{z-1}, \{\tilde{p}_i\}_{i=0}^{z-1})$ -block preserving). Given two orthogonal resolutions of the identity,  $\{p_i\}_{i=0}^{z-1}$  and  $\{\tilde{p}_i\}_{i=0}^{z-1}$ , we say that a quantum channel  $\Phi$  is  $(\{p_i\}_{i=0}^{z-1}, \{\tilde{p}_i\}_{i=0}^{z-1})$ -block preserving if for all  $i, j \in \{0, \dots, z-1\}$ ,

$$\Phi(p_i \mathcal{B}(\mathcal{H}) p_j) \subseteq \tilde{p}_i \mathcal{B}(\mathcal{H}) \tilde{p}_j.$$

From (10.19), if  $\Phi$  is irreducible, then  $\Phi^k$  is  $(\{p_i\}_{i=0}^{z-1}, \{p_{i-k}\}_{i=0}^{z-1})$ -block preserving. Using this notion, the following result shows that PPT channels which are block preserving in the above sense must annihilate off-diagonal blocks. In this work, this is the key property that links PPT channels to eventually entanglement-breaking channels. We note that results of a similar flavor were shown in [Car16; Car19].

**Lemma 11.3.3.** *If  $\Phi$  is a PPT quantum channel and is  $(\{p_i\}_{i=0}^{z-1}, \{\tilde{p}_i\}_{i=0}^{z-1})$ -block preserving, then*

$$\Phi(p_i \mathcal{B}(\mathcal{H}) p_j) = \{0\} \quad (11.6)$$

for all  $i \neq j$ .

*Proof.* Let us prove the contrapositive. Assume for some  $i \neq j$ ,  $\Phi(p_i \mathcal{B}(\mathcal{H}) p_j) \neq \{0\}$ ; without loss of generality, take  $i = 0$  and  $j = 1$ . Then let  $|0\rangle\langle 1| \in p_0 \mathcal{B}(\mathcal{H}) p_1$  such that  $\Phi(|0\rangle\langle 1|) \neq 0$ . Then also  $\Phi(|1\rangle\langle 0|) \neq 0$ . Consider  $\Omega_{01} = |\Omega_{01}\rangle\langle \Omega_{01}|$ , the unnormalized Bell state associated to  $|\Omega_{01}\rangle := |00\rangle + |11\rangle$ . The block-preserving assumption yields

$$\Phi(|i\rangle\langle j|) = \Phi(p_i |i\rangle\langle j| p_j) = \tilde{p}_i \Phi(|i\rangle\langle j|) \tilde{p}_j$$

for each  $i, j \in \{0, 1\}$ . Then, neglecting rows and columns of all zeros,  $(\Phi \otimes \text{id})(\Omega_{01})$  can be written as

$$(\Phi \otimes \text{id})(\Omega_{01}) = \sum_{i,j=0}^1 \Phi(|i\rangle\langle j|) \otimes |i\rangle\langle j| = \left( \begin{array}{cc|cc} \Phi(|0\rangle\langle 0|) & 0 & 0 & \Phi(|1\rangle\langle 0|) \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ \Phi(|0\rangle\langle 1|) & 0 & 0 & \Phi(|1\rangle\langle 1|) \end{array} \right)$$

as  $\{\tilde{p}_0, \tilde{p}_1\}$  blocks in the  $\{|0\rangle, |1\rangle\}$  basis. Now, taking the partial transpose on the first system,

$$(\mathcal{T} \otimes \text{id}) \circ (\Phi \otimes \text{id})(\Omega_{01}) = \left( \begin{array}{cc|cc} \Phi(|0\rangle\langle 0|) & 0 & 0 & 0 \\ 0 & 0 & \Phi(|1\rangle\langle 0|) & 0 \\ \hline 0 & \Phi(|0\rangle\langle 1|) & 0 & 0 \\ 0 & 0 & 0 & \Phi(|1\rangle\langle 1|) \end{array} \right).$$

The eigenvalues of this matrix are the eigenvalues of  $\Phi(|0\rangle\langle 0|)$ , together with the eigenvalues of  $\Phi(|1\rangle\langle 1|)$ , and the eigenvalues of the block matrix

$$X = \begin{pmatrix} 0 & \Phi(|1\rangle\langle 0|) \\ \Phi(|0\rangle\langle 1|) & 0 \end{pmatrix}.$$

Since  $X$  is non-zero, self-adjoint, and traceless, it must have both strictly positive and strictly negative eigenvalues. Thus,  $(\mathcal{T} \otimes \text{id}) \circ (\Phi \otimes \text{id})(\Omega_{01})$  has negative eigenvalues, so  $(\Phi \otimes \text{id})(\Omega_{01})$  is not PPT.  $\square$

**Proposition 11.3.4.** *Let  $\Phi$  be an irreducible CPTP map,  $k \geq 1$ , and let us adopt the notation of Proposition 10.3.8. Assume  $\Phi$  is not primitive (i.e  $z \geq 2$ ). Then*

$$\Phi^k \in \text{PPT}(\mathcal{H}) \implies \Phi^k(p_i \mathcal{B}(\mathcal{H}) p_j) = \{0\} \quad \forall i \neq j. \quad (11.7)$$

On the other hand, if

$$\Phi^k(p_i \mathcal{B}(\mathcal{H}) p_j) = \{0\} \quad \forall i \neq j \quad (11.8)$$

and additionally, for each  $j$  such that  $\text{rank } p_j \geq 2$ ,

$$\|J(\Phi_Q^k|_{p_j \mathcal{B}(\mathcal{H}) p_j})\|_2 \leq z \lambda_{\min}(\sigma|_{p_j \mathcal{H}}) \quad (11.9)$$

then  $\Phi^k \in \text{EB}(\mathcal{H})$ .

In the case  $z = d_{\mathcal{H}}$ , we may write  $p_j = |j\rangle \langle j|$  for  $j = 0, \dots, z-1$ . In this case,  $\Phi_Q^k(|i\rangle \langle j|) = 0$  for all  $i \neq j$ , if and only if  $\Phi^k \in \text{EB}(\mathcal{H})$ .

*Remark.* Under the assumption (11.8),

$$\|J(\Phi_Q^k)\|_2 \leq z \lambda_{\min}(\sigma) \quad (11.10)$$

implies (11.9), as

$$\|J(\Phi_Q^k|_{p_j \mathcal{B}(\mathcal{H}) p_j})\|_2 \leq \sum_{n=0}^{z-1} \|J(\Phi_Q^k|_{p_n \mathcal{B}(\mathcal{H}) p_n})\|_2 = \|J(\Phi_Q^k)\|_2 \leq z \lambda_{\min}(\sigma) \leq z \lambda_{\min}(\sigma|_{p_j}).$$

*Proof.* (11.7) follows immediately from the fact that if  $\Phi$  is irreducible, then  $\Phi^k$  is  $(\{p_i\}_{i=0}^{z-1}, \{p_{i-k}\}_{i=0}^{z-1})$ -block preserving, and Lemma 11.3.3.

Next, assume (11.8) holds. Let  $\{|i\rangle\}_{i=0}^{d_{\mathcal{H}}-1}$  be an orthonormal basis of  $\mathcal{H}$  such that there are disjoint index sets  $\{I_n\}_{n=0}^{z-1}$  such that for each  $n$ ,  $|i\rangle \in p_n \mathcal{H}$  for  $i \in I_n$ . Taking the Choi matrix in the  $|i\rangle$  basis,

$$\begin{aligned} J(\Phi^k) &= \sum_{i,j} \Phi^k(|i\rangle \langle j|) \otimes |i\rangle \langle j| \\ &= \sum_{n=0}^{z-1} \sum_{i,j \in I_n} \Phi^k(|i\rangle \langle j|) \otimes |i\rangle \langle j| \end{aligned}$$

using that  $\Phi^k(|i\rangle\langle j|) = 0$  whenever  $i$  and  $j$  do not share an index set  $I_n$ , which follows from (11.8). Then (see e.g. (10.24)):

$$\begin{aligned} J(\Phi^k) &= \sum_{n=0}^{z-1} \sum_{i,j \in I_n} (\delta_{i,j} z p_{n-k} \sigma + \Phi_Q^k(|i\rangle\langle j|)) \otimes |i\rangle\langle j| \\ &= \sum_{n=0}^{z-1} \left( z p_{n-k} \sigma \otimes p_n + \sum_{i,j \in I_n} \Phi_Q^k(|i\rangle\langle j|) \otimes |i\rangle\langle j| \right). \end{aligned}$$

We note that  $\sum_{i,j \in I_n} \Phi_Q^k(|i\rangle\langle j|) \otimes |i\rangle\langle j| \in p_{n-k} \otimes p_n \mathcal{B}(\mathcal{H} \otimes \mathcal{H}) p_{n-k} \otimes p_n$ , since both  $\Phi$  and  $\Phi_P$  map  $p_n \mathcal{B}(\mathcal{H}) p_n$  to  $p_{n-1} \mathcal{B}(\mathcal{H}) p_{n-1}$  (see eqs. (10.19) and (10.23)). Since by assumption

$$\|J(\Phi_Q^k|_{p_n \mathcal{B}(\mathcal{H}) p_n})\|_2 \leq z \lambda_{\min}(\sigma|_{p_n \mathcal{H}})$$

then Lemma 11.1.1 applied to the Hilbert space  $p_{n-k} \mathcal{H} \otimes p_n \mathcal{H}$  gives that  $z \sigma \otimes \mathbf{1}|_{p_{n-k} \mathcal{H} \otimes p_n \mathcal{H}} + J(\Phi_Q^k|_{p_n \mathcal{B}(\mathcal{H}) p_n})$  is separable on that space. We may embed this state in  $\mathcal{B}(\mathcal{H} \otimes \mathcal{H})$  (without changing the tensor product structure) yielding that

$$p_{n-k} \otimes p_n (z \sigma \otimes \mathbf{1} + J(\Phi_Q^k)) p_{n-k} \otimes p_n$$

is a non-full-rank separable state on  $\mathcal{B}(\mathcal{H} \otimes \mathcal{H})$ . By Lemma 10.2.4, summing over  $n$  yields that  $J(\Phi^k)$  is separable, so  $\Phi^k \in \text{EB}(\mathcal{H})$ .

For the case  $z = d$ , we simply note that  $\text{rank } p_j = 1$  for all  $j$ , and hence the statement follows from the above two results.  $\square$

Since the limit points of  $\{J(\Phi^n)\}_{n=1}^\infty$  are separable but arbitrarily close to entangled states (as shown in the remarks following Proposition 10.3.8) the question arises of whether or not there are quantum channels that are both irreducible and in  $\text{EB}_\infty(\mathcal{H})$ . In the case that  $\Phi$  has maximal period  $z = d_{\mathcal{H}}$ , Lemma 11.3.1 resolves this affirmatively as long as  $\Phi_Q$  is not nilpotent. In the case when the period is much less than the dimension; say  $z = 2 < d_{\mathcal{H}}$ , then the underlying argument (relying on the reshuffling criterion via (11.5)) provides little help:  $\|\Phi^n\|_1 = z + o(n) < d_{\mathcal{H}}$  for large  $n$ . However, using Proposition 11.3.4, we can design  $\text{EB}_\infty(\mathcal{H})$  irreducible channels rather easily, as shown in the following example.

**Example 11.3.5.** Let us construct an irreducible quantum channel  $\Phi$  via Proposition 10.3.8. with period  $z = 2$ . We choose any full rank state  $\sigma$  as the invariant state, and any pair of orthogonal projections  $\{p_1, p_2\}$  which commute with  $\sigma$  as the Perron-Frobenius projections. Let  $d_j = \text{rank } p_j$  for  $j = 0, 1$ , and let  $\{|e_i\rangle\}_{i=0}^{d_0-1}$  be an eigenbasis of  $\sigma p_0$  and

likewise  $\{|f_i\rangle\}_{i=0}^{d_1-1}$  be an eigenbasis of  $\sigma p_1$ . For some  $\lambda \in \mathbb{C}$  with  $|\lambda| < 1$ , define  $\Phi_Q$  by

$$\begin{aligned}\Phi_Q(|e_i\rangle\langle e_j|) &= \Phi_Q(|f_i\rangle\langle f_j|) = 0, \\ \Phi_Q(|e_i\rangle\langle f_j|)^* &= \Phi_Q(|f_i\rangle\langle e_j|) = \delta_{i,0}\delta_{j,0}\bar{\lambda}|e_0\rangle\langle f_0|\end{aligned}$$

for each  $i, j$ . Then  $\Phi_Q(\sigma p_j) = \Phi_Q^*(p_j) = 0$ ,  $\text{spr}(\Phi_Q) = |\lambda| < 1$ . Since

$$J(\Phi_Q) = \lambda|f_0\rangle\langle e_0| \otimes |f_0\rangle\langle e_0| + \bar{\lambda}|e_0\rangle\langle f_0| \otimes |f_0\rangle\langle e_0|,$$

we have  $\|J(\Phi_Q)\|_2 = \sqrt{2}|\lambda|$ . Thus, choosing  $\lambda$  to satisfy  $0 < |\lambda| < \min(1, \sqrt{2}\lambda_{\min}(\sigma))$ , we have  $J(\Phi_Q) \geq -J(\Phi_P)$  by (10.26), and  $\Phi$  is an irreducible CPTP map of period 2. Moreover,

$$\Phi_Q^n(|e_0\rangle\langle f_0|) = \begin{cases} \lambda^n |f_0\rangle\langle e_0| & n \text{ odd} \\ \lambda^n |e_0\rangle\langle f_0| & n \text{ even} \end{cases}$$

which is non-zero for any  $n$ . Thus,  $\Phi \in \text{EB}_\infty(\mathcal{H})$  by Proposition 11.3.4.

*Remark.* It was proven in Theorem 21 of [LG16] that if the number of zero eigenvalues of  $\Phi$  is strictly less than  $2(d_{\mathcal{H}} - 1)$ , the fact that  $\Phi$  has at least two peripheral eigenvalues implies that it is entanglement saving. However, in the present example,  $\Phi$  has four non-zero eigenvalues (two peripheral, and  $\pm|\lambda|$ ), and therefore  $d_{\mathcal{H}}^2 - 4$  zero eigenvalues. Thus, for  $d_{\mathcal{H}} \geq 3$ , Theorem 11 of [LG16] does not apply to  $\Phi$ .

Proposition 11.3.4 implies the following corollary, which is extended to the non-irreducible case in the next section.

**Corollary 11.3.6.** *Any irreducible PPT channel  $\Phi : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H})$  is eventually entanglement breaking.*

*Proof.* Since the channel is PPT, (11.8) holds. Then setting  $\ell := \text{spr}(\Phi_Q) < 1$ , by Gelfand's formula we have that for any  $\varepsilon \in (0, 1 - \ell)$ , there exists  $n_0 > 0$  such that for all  $k \geq n_0$ ,

$$\|J(\Phi_Q^k)\|_2 = \|\Phi_Q^k\|_2 \leq (\ell + \varepsilon)^k.$$

Thus, for  $k$  large enough,  $\|J(\Phi_Q^k)\|_2 \leq z\lambda_{\min}(\sigma)$  and (11.9) holds. Hence,  $\Phi^k \in \text{EB}(\mathcal{H})$ .  $\square$

*Remark.* [HRS20] provides many quantitative estimates of entanglement-breaking times which are omitted here.

### 11.3.2 Beyond irreducibility

A non-irreducible channel can be decomposed into *irreducible components*, on which it acts irreducibly. More specifically, we may decompose the identity  $\mathbf{1}$  of  $\mathcal{H}$  into maximal subspaces with corresponding orthogonal projections  $D, P_1, \dots, P_n$  such that  $\Phi$  restricted to  $P_i\mathcal{B}(\mathcal{H})P_i$  is irreducible, and  $D\mathcal{H}$  is orthogonal to the support of every invariant state of  $\Phi$  [CP16]. In particular,  $D = 0$  is equivalent to  $\Phi$  being a faithful quantum channel (that is, possessing a full-rank invariant state). In general,  $\Phi$  may act non-trivially on  $P_i\mathcal{B}(\mathcal{H})P_j$  for  $i \neq j$ . The following proposition shows that this is not the case for PPT maps  $\Phi$ .

**Theorem 11.3.7.** *Any faithful PPT channel  $\Phi : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H})$  is the direct sum of irreducible PPT quantum channels, and therefore is eventually entanglement breaking.*

*Remark.* This result extends that of [KMP17] where it was shown that PPT maps are AEB. It also completes Theorem 4.4 of [RJP18] where it was shown that any unital PPT channel is EEB.

*Proof.* As  $\Phi$  has an invariant state of full rank, it follows from [CP16] that we may decompose the identity  $\mathbf{1}$  of  $\mathcal{H}$  into maximal subspaces with corresponding orthogonal projections  $P_1, \dots, P_n$  s.t.  $\Phi$  restricted to  $P_i\mathcal{B}(\mathcal{H})P_i$  is irreducible. We will call  $P_i\mathcal{B}(\mathcal{H})P_i$  a maximal irreducible component. From [CP16, Proposition 5.4], we have that

$$\Phi(P_i\mathcal{D}(\mathcal{H})P_j) \subseteq P_i\mathcal{B}(\mathcal{H})P_j$$

and hence linearity yields

$$\Phi(P_i\mathcal{B}(\mathcal{H})P_j) \subseteq P_i\mathcal{B}(\mathcal{H})P_j.$$

Thus,  $\Phi$  is  $(\{P_i\}_{i=1}^n, \{P_i\}_{i=1}^n)$ -block preserving, and by Lemma 11.3.3,

$$\Phi(P_i\mathcal{B}(\mathcal{H})P_j) = \{0\}$$

for each  $i \neq j$ . Hence,  $\Phi = \bigoplus_{i=1}^n \Phi_i$  for  $\Phi_i = \Phi|_{P_i\mathcal{B}(\mathcal{H})P_i}$ . Each  $\Phi_i$  is irreducible and PPT and hence EEB by Corollary 11.3.6. Thus,  $\Phi$  is EEB as well.  $\square$

Combining Theorem 11.3.7 with that observation that for an irreducible channel  $\Phi$  of period  $z$ , the channel  $\Phi^z$  is the direct sum of primitive channels leads to the following structural result.

**Theorem 11.3.8.** *Let  $\Phi : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H})$  be a faithful quantum channel. The following are equivalent.*

1.  $\Phi \in \text{EEB}(\mathcal{H})$ ,
2.  $\Phi$  is eventually PPT, meaning  $\Phi^n \in \text{PPT}(\mathcal{H})$  for some  $n \in \mathbb{N}$ ,
3.  $\Phi^{d_{\mathcal{H}}^2 Z}$  is the direct sum of primitive channels, where  $Z$  is the least common multiple of the periods of  $\Phi$  restricted to each of its irreducible components,
4.  $\Phi^n$  is the direct sum of primitive channels for some  $n \leq d_{\mathcal{H}}^2 \exp(1.04 d_{\mathcal{H}})$ .

*Remark.* This provides a characterization of  $\text{EEB}(\mathcal{H})$  for faithful channels.

*Proof.* We will show the chain of implications

$$(4) \implies (1) \implies (2) \implies (3) \implies (4),$$

establishing the stated equivalence.

Since primitive channels are eventually entanglement breaking<sup>2</sup>, the direct sum of primitive channels  $\Phi = \bigoplus_i \Phi_i$  is also eventually entanglement breaking, as

$$J(\Phi^n) = \bigoplus_i J(\Phi_i^n)$$

is separable once each  $J(\Phi_i^n)$  is separable. Thus, (4) implies (1).

Next, that (1) implies (2) is immediate. Let us see that (2) implies (3). Let  $\Phi \in \text{EEB}(\mathcal{H})$  be a faithful quantum channel which is eventually PPT. Without loss of generality, assume  $\Phi$  is not irreducible. Recall from the proof of Theorem 11.3.7 that there exist  $P_1, \dots, P_k \in \mathcal{B}(\mathcal{H})$  distinct orthogonal projections such that  $\Phi(P_i \mathcal{B}(\mathcal{H}) P_i) \subseteq \Phi(P_i \mathcal{B}(\mathcal{H}) P_i)$  with  $\sum_i P_i = \mathbf{1}$ , which moreover satisfy

$$\Phi(P_i \mathcal{B}(\mathcal{H}) P_j) \subseteq \Phi(P_i \mathcal{B}(\mathcal{H}) P_j)$$

for all  $i$  and  $j$ , and  $\Phi_i = \Phi|_{P_i \mathcal{B}(\mathcal{H}) P_i}$  is irreducible, with some period  $z_i$ , and Perron-Frobenius projections  $\{p_k^{(i)}\}_{k=0}^{z_i-1}$ . Let  $Z$  be the least common multiple of the periods  $z_i$ . Then

$$\Phi^Z(p_i^{(j)} P_j \mathcal{B}(\mathcal{H}) P_k p_\ell^{(k)}) \subseteq p_i^{(j)} P_j \mathcal{B}(\mathcal{H}) P_k p_\ell^{(k)} \quad (11.11)$$

for each  $i, j, k, \ell$ . Moreover, since  $\Phi$  is eventually PPT,  $\Phi^{nZ}$  is PPT for some  $n \in \mathbb{N}$ . Hence, by Lemma 11.3.3 and (11.11),

$$\Phi^{nZ}(p_i^{(j)} P_j \mathcal{B}(\mathcal{H}) P_k p_\ell^{(k)}) = \{0\}$$

---

<sup>2</sup>As shown in Section 11.1



unless  $j = k$  and  $i = \ell$ . For  $j \neq k$  and  $i \neq \ell$ ,  $\Phi^Z|_{p_i^{(j)}P_j\mathcal{B}(\mathcal{H})P_kp_\ell^{(k)}}$  is therefore a nilpotent linear endomorphism (on the vector space  $p_i^{(j)}P_j\mathcal{B}(\mathcal{H})P_kp_\ell^{(k)}$ ). Since  $\dim(p_i^{(j)}P_j\mathcal{B}(\mathcal{H})P_kp_\ell^{(k)}) \leq d_{\mathcal{H}}^2$ , we have

$$\Phi^{d^2Z}(p_i^{(j)}P_j\mathcal{B}(\mathcal{H})P_kp_\ell^{(k)}) = \{0\}$$

Hence,  $\Phi^{d^2Z} = \bigoplus_{i=1}^n \bigoplus_{j=0}^{z_i-1} \Phi_{ij}$  is the direct sum of primitive quantum channels  $\Phi_{ij} := \Phi^{d^2Z}|_{\mathcal{B}(p_jP_i\mathcal{H})}$ .

To prove (3) implies (4), we upper bound  $Z$ . Since each period is at most  $d_{\mathcal{H}}$ ,  $Z$  is at most the least common multiple of the natural numbers  $1, \dots, d_{\mathcal{H}}$ . By [RS62, Theorem 12], we thus have

$$Z \leq \exp(1.03883 d_{\mathcal{H}}).$$

□

### 11.3.3 Reduction to the trace-preserving case

When considering quantum channels as a form of time evolution, trace-preservation is a natural assumption as an analog of conserving total probability. When considering maps resulting from the inverse Choi isomorphism  $J^{-1}$  applied to bipartite operators, however, the map  $J^{-1}(\rho_{A\bar{A}})$  is trace-preserving if and only if the first marginal is completely mixed:  $\rho_A = \mathbf{1}_A$ . Thus, a priori, results about quantum channels only provide results about a restricted class of bipartite operators. However, as shown by Proposition 11.3.9 below, a large class of CP maps are (up to normalization) similar to CPTP maps. Moreover, this similarity transformation preserves many properties.

Recall that  $\text{spr}(\Phi)$  denotes the *spectral radius* of a map  $\Phi$ , defined as  $\text{spr}(\Phi) := \max_{\lambda} |\lambda|$ , where  $\lambda$  ranges over the eigenvalues of  $\Phi$ , and that the spectral radius of a quantum channel is 1.

**Proposition 11.3.9.** *Let  $\Phi$  be a CP map such that for some  $X > 0$ ,  $\Phi^*(X) = \text{spr}(\Phi)X$  and for some  $\sigma > 0$ ,  $\Phi(\sigma) = \text{spr}(\Phi)\sigma$ . Then*

$$\mathcal{P}_X(\Phi) := \frac{1}{\text{spr}(\Phi)} \Gamma_X \circ \Phi \circ \Gamma_X^{-1}, \quad \text{where } \Gamma_X(Y) := X^{1/2}YX^{1/2} \quad (11.12)$$

has the following properties:

1.  $\mathcal{P}_X(\Phi)$  is a faithful quantum channel
2.  $\mathcal{P}_X(\Phi^n) = \mathcal{P}_X(\Phi)^n$ , where powers denote repeated composition
3.  $\mathcal{P}_X(\Phi)$  is EB if and only if  $\Phi$  is EB

4.  $\mathcal{P}_X(\Phi)$  is PPT if and only if  $\Phi$  is PPT

*Remarks.*

- This transformation is not novel; the map  $\Gamma_\sigma \circ \Phi^* \circ \Gamma_{\Phi(\sigma)}^{-1}$  is known as the Petz recovery map [Pet03], and is usually considered in the case of CPTP maps  $\Phi$ . Replacing  $\Phi$  with  $\Phi^*$  and  $\sigma$  with  $X$  yields (11.12) in the case that  $\Phi^*(X) = \text{spr}(\Phi)X$ . This transformation has also been considered in [HJPR18, Appendix A] in order to study the spectral properties of deformed CP maps as a function of the deformation. See [Wol12, Theorem 3.2] for a similar but slightly different approach (using  $\Phi^*(\mathbf{1})$  instead of  $X$ ).
- By the Perron-Frobenius theory (see [EH78, Theorem 2.5]), if  $\Phi$  is a positive map, then the spectral radius  $\text{spr}(\Phi)$  is an eigenvalue of  $\Phi$ , and  $\Phi$  admits a positive semidefinite eigenvector  $\sigma$ , which can be normalized to have unit trace. Since  $\text{spr}(\Phi) = \text{spr}(\Phi^*)$ , the same logic applied to  $\Phi^*$  yields  $X \geq 0$  such that  $\Phi^*(X) = \text{spr}(\Phi)X$ . Hence, the assumption of Proposition 11.3.9 is that both of these eigenvectors are full rank.
- This transformation cannot be applied in general to a CPTP map  $\Phi$  in order to obtain a CPTP unital map  $\tilde{\Phi}$ , since trace-preservation will be lost. An intuitive way to see this is that a similarity transformation corresponds to a choice of (non-orthogonal) basis, and by fixing  $\Phi^*(\mathbf{1}) = \mathbf{1}$ , we choose a particular basis for  $\Phi$  in which the dual eigenvector for  $\text{spr}(\Phi)$  is represented by the identity matrix. Thus, in general, we cannot simultaneously choose a basis to fix a representation for the eigenvector of  $\text{spr}(\Phi)$ .

*Proof.* 1.  $\mathcal{P}_X(\Phi)$  is the positive multiple of the composition of CP maps and hence is CP. Since

$$\mathcal{P}_X(\Phi)^* = \frac{1}{\text{spr}(\Phi)} \Gamma_X^{-1} \circ \Phi^* \circ \Gamma_X$$

we have that  $\mathcal{P}_X(\Phi)^*$  is unital:

$$\mathcal{P}_X(\Phi)^*(\mathbf{1}) = \frac{1}{\text{spr}(\Phi)} \Gamma_X^{-1} \circ \Phi^*(X) = \Gamma_X^{-1}(X) = \mathbf{1}$$

and hence  $\mathcal{P}_X(\Phi)$  is TP. Lastly,  $\rho := \frac{\Gamma_X(\sigma)}{\text{tr}[\dots]}$  is a full-rank invariant state for  $\mathcal{P}_X(\Phi)$ : up to normalization,

$$\mathcal{P}_X(\Phi)(\rho) = \frac{1}{\text{spr}(\Phi)} \Gamma_X \circ \Phi \circ \Gamma_X^{-1} \circ \Gamma_X(\sigma) = \frac{1}{\text{spr}(\Phi)} \Gamma_X \circ \Phi(\sigma) = \Gamma_X(\sigma) = \rho$$

and hence  $\mathcal{P}_X(\Phi)$  is faithful.

2. We have

$$\begin{aligned}
 \mathcal{P}_X(\Phi^n) &= \frac{1}{\text{spr}(\Phi^n)} \Gamma_X \circ \Phi^n \circ \Gamma_X^{-1} \\
 &= \frac{1}{\text{spr}(\Phi)^n} \Gamma_X \circ \Phi^{n-1} \circ \Gamma_X^{-1} \circ \Gamma_X \circ \Phi \circ \Gamma_X^{-1} \\
 &= \dots \\
 &= \frac{1}{\text{spr}(\Phi)^n} \Gamma_X \circ \Phi \circ \Gamma_X^{-1} \circ \Gamma_X \circ \Phi \circ \Gamma_X^{-1} \circ \dots \circ \Gamma_X \circ \Phi \circ \Gamma_X^{-1} \\
 &= \mathcal{P}_X(\Phi)^n.
 \end{aligned}$$

3. If  $\Phi$  is EB, then  $\Phi \otimes \text{id}(\mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)_+) \subseteq \text{SEP}(A : B)$ . Since  $X$  is invertible,  $\Gamma_X^{-1} \otimes \text{id}(\mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)_+) = \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)_+$ . Hence  $\Phi \circ \Gamma_X^{-1} \otimes \text{id}$  is EB. Since SEP is invariant under local positive operations,  $\mathcal{P}_X(\Phi)$  is EB too. The converse follows the same way.
4. The same proof holds with  $\text{SEP}(A : B)$  replaced by  $\text{PPT}(A : B)$  and EB replaced by PPT.

□

This transformation allows us to establish our most general result on the entanglement-breakability of PPT maps.

**Theorem 11.3.10.** *Let  $\Phi$  be a CP map such that for some  $X > 0$ ,  $\Phi^*(X) = \text{spr}(\Phi)X$  and for some  $\sigma > 0$ ,  $\Phi(\sigma) = \text{spr}(\Phi)\sigma$ . Then if  $\Phi$  is PPT, it is EEB.*

*Proof.* By (1),  $\mathcal{P}_X(\Phi)$  is faithful, CPTP, and by (4) PPT. Hence, for some  $n$ ,  $\mathcal{P}_X(\Phi)^n = \mathcal{P}_X(\Phi^n)$  is EB, using (2) and Theorem 11.3.7. Then  $\Phi^n$  is EB by (3). □



# Chapter 12

## Repeated interaction systems

In the previous chapter, we investigated the finite-time entanglement breaking properties of discrete quantum Markov semigroups (QMS)  $\{\Phi^n\}_{n=1}^\infty$ . In this chapter, we will discuss repeated interaction systems, which provide a physical model from which such QMS, and others, arise. We will see how delving into more of the details of the physical model permits us to make a more refined analysis of the evolution of an initial state. In particular, we will use a two-time measurement protocol in which *indirect* measurements are performed on the system. The state of the system and other quantities of interest can then be seen as random variables with respect to the distribution of measurement outcomes. We will chiefly investigate the so-called *entropy production* of the evolution, however, rather than entanglement-breaking properties of the maps in question.

We will also consider a generalization of the setting of the previous chapter. In the systems considered in that chapter, the evolution of a state after  $n$  timesteps is given by  $\underbrace{\Phi \circ \Phi \circ \dots \circ \Phi}_{n \text{ times}}$ . In this chapter, we consider systems in which the evolution over  $n$  timesteps is given by a sequence  $\Phi_n \circ \Phi_{n-1} \circ \dots \circ \Phi_1$  where one may have  $\Phi_1 \neq \Phi_2$  and so forth. However, the maps  $\{\Phi_k\}_k$  will be obtained in such a way as to guarantee that consecutive maps differ only by a small amount  $O(1/T)$  when the total number of timesteps is given by  $T$ . This is the so-called adiabatic setting discussed in Section 12.5.

This chapter discusses the results of [HJPR18] with the aim of illustrating how the use of the two-time measurement protocol (defined in Section 12.2) yields more detailed information about the evolution of the system and its statistics. For the sake of focus and brevity, many of the proofs of the technical results have been suppressed (but may be found in the full text of [HJPR18]).

## 12.1 Repeated interaction systems

A repeated interaction system (RIS) consists of the system of interest  $\mathcal{S}$ , along with a structured environment  $\mathcal{E}$  which consists of a chain of “probes”  $\mathcal{E}_k$ ,  $k = 1, 2, \dots$ . The system  $\mathcal{S}$  interacts with  $\mathcal{E}_k$  (and only  $\mathcal{E}_k$ ) during the time interval  $[k\tau, (k+1)\tau)$ , where  $\tau > 0$  is the interaction time.

More specifically, we describe the quantum system  $\mathcal{S}$  by a finite-dimensional Hilbert space  $\mathcal{H}_{\mathcal{S}}$ , a (time-independent) Hamiltonian  $h_{\mathcal{S}} = h_{\mathcal{S}}^* \in \mathcal{B}(\mathcal{H}_{\mathcal{S}})$ , and an initial state  $\rho^i \in \mathcal{D}(\mathcal{H}_{\mathcal{S}})$ . Likewise, the  $k$ th quantum probe  $\mathcal{E}_k$  is described by a finite dimensional Hilbert space  $\mathcal{H}_{\mathcal{E},k}$ , Hamiltonian  $h_{\mathcal{E}_k} = h_{\mathcal{E}_k}^* \in \mathcal{B}(\mathcal{H}_{\mathcal{E},k})$ , and initial state  $\xi_k^i \in \mathcal{D}(\mathcal{H}_{\mathcal{E},k})$ . We assume the probe Hilbert spaces  $\mathcal{H}_{\mathcal{E},k}$  are all identical,  $\mathcal{H}_{\mathcal{E},k} \equiv \mathcal{H}_{\mathcal{E}}$ , and that the initial state of each probe is a Gibbs state at inverse temperature  $\beta_k > 0$ :

$$\xi_k^i = \frac{e^{-\beta_k h_{\mathcal{E}_k}}}{\text{tr}(e^{-\beta_k h_{\mathcal{E}_k}})}.$$

We often use  $Z_{\beta,k}$  to denote the partition function  $\text{tr}(e^{-\beta_k h_{\mathcal{E}_k}})$ . We will consider a finite number  $T \in \mathbb{N}$  of probes as well as the limit  $T \rightarrow \infty$ .

The state of the system  $\mathcal{S}$  evolves by interacting with each probe, one at a time, as follows. Assume that after interacting with the first  $k-1$  probes the state of the system is  $\rho_{k-1}$ . Then the system and the  $k$ th probe, with joint initial state  $\rho_{k-1} \otimes \xi_k^i$ , evolve for a time  $\tau$  via the free Hamiltonian plus interaction  $v_k$  according to the unitary operator

$$U_k := \exp(-i\tau(h_{\mathcal{S}} \otimes \text{id} + \text{id} \otimes h_{\mathcal{E}_k} + v_k)),$$

yielding a joint final state  $U_k(\rho_{k-1} \otimes \xi_k^i)U_k^*$ . The probe  $\mathcal{E}_k$  is traced out, resulting in the system state

$$\rho_k := \text{tr}_{\mathcal{E}}(U_k(\rho_{k-1} \otimes \xi_k^i)U_k^*),$$

where  $\text{tr}_{\mathcal{E}}$  is the partial trace over  $\mathcal{H}_{\mathcal{E}}$ ; likewise,  $\text{tr}_{\mathcal{S}}$  denotes the partial trace over  $\mathcal{H}_{\mathcal{S}}$ . We also introduce  $\xi_k^f := \text{tr}_{\mathcal{S}}(U_k(\rho_{k-1} \otimes \xi_k^i)U_k^*)$ . A diagram representing a repeated interaction system is shown in Figure 12.1.

The evolution of the system  $\mathcal{S}$  during the  $k$ th step is given by the *reduced dynamics*

$$\begin{aligned} \mathcal{L}_k : \mathcal{B}(\mathcal{H}_{\mathcal{S}}) &\rightarrow \mathcal{B}(\mathcal{H}_{\mathcal{S}}) \\ \eta &\mapsto \text{tr}_{\mathcal{E}}(U_k(\eta \otimes \xi_k^i)U_k^*), \end{aligned} \tag{12.1}$$

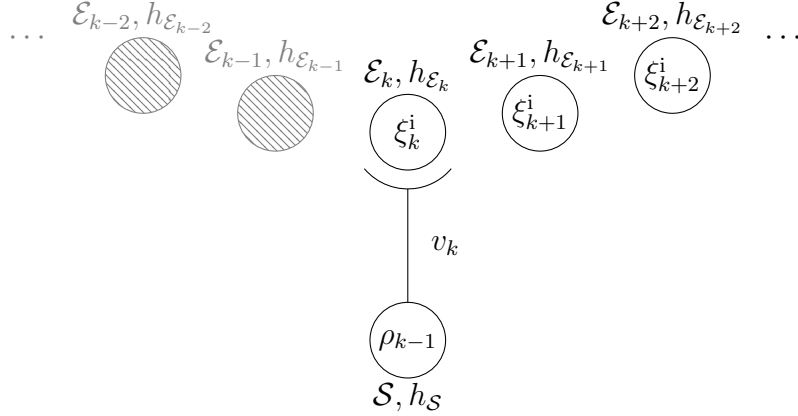


Fig. 12.1 A schematic representation of a repeated interaction system, in which the system  $\mathcal{S}$  interacts, one at a time, with a chain of probes  $\mathcal{E}_1, \mathcal{E}_2, \dots$ , at step  $k$  of the interaction. The  $k$ th probe has a Hamiltonian  $h_{\mathcal{E}_k}$  and initial state  $\xi_k^i$ , and interacts with the system via a coupling  $v_k$ . Before the  $k$ th interaction, the system, which has its own Hamiltonian  $h_{\mathcal{S}}$ , starts in state  $\rho_{k-1}$ .

that is  $\rho_k = \mathcal{L}_k(\rho_{k-1})$ . The reduced dynamics  $\mathcal{L}_k$  is completely positive and trace-preserving, and hence maps  $\mathcal{D}(\mathcal{H}_{\mathcal{S}})$  to  $\mathcal{D}(\mathcal{H}_{\mathcal{S}})$ . By iterating this evolution, we find that the state of the system  $\mathcal{S}$  after  $k$  steps is given by the composition

$$\rho_k = (\mathcal{L}_k \circ \dots \circ \mathcal{L}_1)(\rho^i). \quad (12.2)$$

We often omit the parentheses and composition symbols. In the case that the probes have identical Hamiltonians,  $h_{\mathcal{E}} \equiv h_{\mathcal{E}_k}$ , and identical temperatures  $\beta \equiv \beta_k$ , and the system and each probe interacts with the system via identical couplings  $v \equiv v_k$ , one obtains identical reduced dynamics  $\mathcal{L} \equiv \mathcal{L}_k$ . In this case,  $\rho_k = \mathcal{L}^k(\rho^i)$ , where  $\mathcal{L}^k$  denotes the  $k$ -fold composition of  $\mathcal{L}$  with itself. Hence, in this case the evolution of the system forms a discrete-time QMS as discussed in the previous chapter. On the other hand, the Stinespring dilation shows that any discrete-time QMS can be formulated as a repeated interaction system with reduced dynamics  $\mathcal{L} \equiv \mathcal{L}_k$  (albeit one in which the probe has zero temperature, i.e.  $\beta \equiv \beta_k = +\infty$ ).

## 12.2 Two-time measurement protocols

We now describe a two-time measurement protocol for repeated interaction systems with  $T$  probes. By studying quantities of interest as random variables on the distribution of measurement outcomes, we obtain the so-called *full statistics*. We will see that in

many cases this provides a generalization of the setting in which measurements are not performed, as the analogous quantities in the non-measured setting can be obtained as expectation values of the random variables obtained in the measured setting.

Note that a similar protocol was considered in [HP13] (see also [BJPP18]). For the purpose of defining the full statistics measure for an RIS, we will consider observables to be measured on both the system  $\mathcal{S}$  and the probes  $\mathcal{E}_k$ ,  $k \in \mathbb{N}$ .

First, we assume that we are given two observables  $A^i$  and  $A^f$  in  $\mathcal{B}(\mathcal{H}_{\mathcal{S}})$  with spectral decompositions

$$A^i = \sum_{a^i} a^i \pi_{a^i}^i, \quad A^f = \sum_{a^f} a^f \pi_{a^f}^f$$

where  $a^i, a^f$  run over the distinct eigenvalues of  $A^i, A^f$  respectively, and  $\pi_{a^i}^i, \pi_{a^f}^f$  denote the corresponding spectral projectors. When we consider increasing the number of probes  $T$ , we assume that the observable  $A^i$  is independent of  $T$  (as we measure it on  $\mathcal{S}$  before the system interacts with any number of probes), but allow  $A^f$  to depend on  $T$ , as long as the family  $(A^f)_{T=1}^\infty$  is uniformly bounded in  $T$ .

On the chain, we consider probe observables  $Y_k \in \mathcal{B}(\mathcal{H}_{\mathcal{E}})$  to be measured on the probe  $\mathcal{E}_k$ . We require that each observable commutes with the corresponding probe Hamiltonian:

$$[Y_k, h_{\mathcal{E}_k}] = 0.$$

We write the spectral decomposition of each  $Y_k$  as

$$Y_k = \sum_{i_k} y_{i_k} \Pi_{i_k}^{(k)}.$$

If the  $k$ th probe is initially in the state  $\xi$ , a measurement of  $Y_k$  before the time evolution will yield  $y_{i_k}$  with probability  $\text{tr}(\xi \Pi_{i_k}^{(k)})$ .

Associated to the observables  $A^i, A^f$  and  $(Y_k)_{k=1}^T$  and the state  $\rho^i$ , we define two processes: the *forward process*, and the *backward process*.

**The forward process** The system  $\mathcal{S}$  starts in some initial state  $\rho^i \in \mathcal{D}(\mathcal{H}_{\mathcal{S}})$  and the probe  $\mathcal{E}_k$  starts in the initial Gibbs state  $\xi_k \in \mathcal{D}(\mathcal{H}_{\mathcal{E}_k})$ ; we write the state of the chain of  $T$  probes  $\Xi = \bigotimes_{k=1}^T \xi_k$ . We measure  $A^i$  on  $\mathcal{S}$  and measure  $Y_k$  on  $\mathcal{E}_k$  for each  $k = 1, \dots, T$ . We obtain results  $a^i$  and  $\vec{i} = (i_k)_{k=1}^T$  with probability

$$\text{tr}((\rho^i \otimes \Xi)(\pi_{a^i}^i \otimes \Pi_{\vec{i}})),$$



where  $\Pi_{\vec{i}} := \bigotimes_{k=1}^T \Pi_{i_k}^{(k)}$ . Then the system interacts with each probe, one at a time, starting at  $k = 1$  until  $k = T$ , via the time evolution

$$U_k := \exp \left( -i\tau(h_{\mathcal{S}} + h_{\mathcal{E}_k} + v_k) \right).$$

Next, we measure  $A^f$  on the system and measure  $Y_k$  on  $\mathcal{E}_k$  for each  $k = 1, \dots, T$ , yielding outcomes  $a^f$  and  $\vec{j} = (j_k)_{k=1}^T$ . The probability of measuring the sequence  $(a^i, a^f, \vec{i}, \vec{j})$  of outcomes is given by

$$\text{tr} \left( U_T \cdots U_1 (\pi_{a^i}^i \otimes \Pi_{\vec{i}}) (\rho^i \otimes \Xi) (\pi_{a^i}^i \otimes \Pi_{\vec{i}}) U_1^* \cdots U_T^* (\pi_{a^f}^f \otimes \Pi_{\vec{j}}) \right).$$

We emphasize that the outcomes are labelled by  $(a^i, a^f, \vec{i}, \vec{j})$  which refers to the eigenprojectors of the operators involved, but not to the corresponding eigenvalues which only need to be distinct. Also, we may write the second measurement projector  $\pi_{a^f}^f \otimes \Pi_{\vec{j}}$  only once by cyclicity of the trace.

**The backward process** The system starts in the state

$$\rho_T^f := \text{tr}_{\mathcal{E}} \left( U_T \cdots U_1 (\rho^i \otimes \Xi) U_1^* \cdots U_T^* \right),$$

and the probe  $\mathcal{E}_k$  starts in the state  $\xi_k$ . We measure observable  $A^f$  on  $\mathcal{S}$  and  $Y_k$  on  $\mathcal{E}_k$  for each  $k = 1, \dots, T$ , yielding outcomes  $a^f$  and  $(j_k)_{k=1}^T$ . Then the system interacts with each probe, one at a time, starting with  $k = T$  until  $k = 1$ , via the time evolution

$$U_k^* = \exp \left( i\tau(h_{\mathcal{S}} + h_{\mathcal{E}_k} + v_k) \right).$$

Then we measure  $A^i$  on  $\mathcal{S}$  and  $Y_k$  on  $\mathcal{E}_k$  for each  $k = 1, \dots, T$ , yielding outcomes  $a^i$  and  $(i_k)_{k=1}^T$ . The probability of these outcomes is given by

$$\text{tr} \left( U_1^* \cdots U_T^* (\pi_{a^f}^f \otimes \Pi_{\vec{j}}) (\rho_T^f \otimes \Xi) (\pi_{a^f}^f \otimes \Pi_{\vec{j}}) U_T \cdots U_1 (\pi_{a^i}^i \otimes \Pi_{\vec{i}}) \right).$$

**The full statistics associated to the two-step measurement process** For notational simplicity, we assume that the cardinality of  $\text{spec } Y_k$  does not depend on  $k$ . We can therefore use the same index set  $\mathfrak{J}$  for all eigenvalue sets:  $\text{spec } Y_k = (y_{i_k})_{i_k \in \mathfrak{J}}$  for all  $k = 1, \dots, T$ . We define the space

$$\Omega_T := \text{spec } A^i \times \text{spec } A^f \times \mathfrak{J}^T \times \mathfrak{J}^T$$

and equip it with the maximal  $\sigma$ -algebra  $\mathcal{P}(\Omega_T)$ . We will refer to elements  $(a^i, a^f, \vec{i}, \vec{j})$  of  $\Omega_T$  as *trajectories*, and denote them by the letter  $\omega$ .

**Definition 12.2.1.** On  $\Omega_T$ , we call the law of the outcomes for the forward process,

$$\mathbb{P}_T^F(a^i, a^f, \vec{i}, \vec{j}) := \text{tr} \left( U_T \cdots U_1 (\pi_{a^i}^i \otimes \Pi_{\vec{i}}) (\rho^i \otimes \Xi) (\pi_{a^i}^i \otimes \Pi_{\vec{i}}) U_1^* \cdots U_T^* (\pi_{a^f}^f \otimes \Pi_{\vec{j}}) \right), \quad (12.3)$$

the forward *full statistics measure*. We denote by  $\mathbb{E}_T$  the expectation with respect to  $\mathbb{P}_T^F$ . We also consider the backward full statistics measure

$$\mathbb{P}_T^B(a^i, a^f, \vec{i}, \vec{j}) := \text{tr} \left( U_1^* \cdots U_T^* (\pi_{a^f}^f \otimes \Pi_{\vec{j}}) (\rho^f \otimes \Xi) (\pi_{a^f}^f \otimes \Pi_{\vec{j}}) U_T \cdots U_1 (\pi_{a^i}^i \otimes \Pi_{\vec{i}}) \right) \quad (12.4)$$

which is the law of the outcomes for the backward process. Let us emphasize here that  $\mathbb{P}_T^F$  and  $\mathbb{P}_T^B$  depend on the spectral projectors  $(\pi_{a^i}^i)_{a^i}$  of  $A^i$ ,  $(\pi_{a^f}^f)_{a^f}$  of  $A^f$ , and  $(\Pi_{\vec{i}})$  of the  $(Y_k)_k$ , and not on the spectral values of these operators. In particular, the probabilities  $\mathbb{P}_T^F$  and  $\mathbb{P}_T^B$  associated with two families of observables  $(Y(s))_{s \in [0,1]}$ ,  $(\tilde{Y}(s))_{s \in [0,1]}$  that have the same spectral projectors (as e.g.  $Y(s) = \beta(s)h_{\mathcal{E}}(s)$  and  $\tilde{Y}(s) = h_{\mathcal{E}}(s)$ ) will be the same.

To  $(Y_k)_{k=1}^T$ ,  $A^i$ , and  $A^f$ , we associate two generic classical random variables on  $(\Omega_T, \mathcal{P}(\Omega_T))$ :

$$\Delta a_T(a^i, a^f, \vec{i}, \vec{j}) := a^i - a^f, \quad (12.5)$$

$$\Delta y_T^{\text{tot}}(a^i, a^f, \vec{i}, \vec{j}) := \sum_{k=1}^T (y_{j_k}^{(k)} - y_{i_k}^{(k)}). \quad (12.6)$$

Note that the assumption that  $(A^f)_{T=1}^\infty$  has uniformly bounded norm yields that the random variable  $\Delta a_T$  has  $L^\infty$  norm uniformly bounded in  $T$ .

## 12.3 Landauer's Principle and the balance equation

Next, we define the quantities of interest whose full statistics we will subsequently study. At this stage, we are studying the repeated interaction system without performing any measurements. For each step  $k$  of the RIS process, we define the quantities

$$\begin{aligned} \Delta S_k &:= S(\rho_{k-1}) - S(\rho_k), \\ \Delta Q_k &:= \text{tr}_{\mathcal{E}}(h_{\mathcal{E}_k} \xi_k^f) - \text{tr}_{\mathcal{E}}(h_{\mathcal{E}_k} \xi_k^i), \end{aligned}$$

that represent the decrease in entropy of the small system, and the increase in energy of probe  $k$ , respectively, and

$$\sigma_k := D(U_k(\rho_{k-1} \otimes \xi_k^i)U_k^* \| \rho_k \otimes \xi_k^i), \quad (12.7)$$

the *entropy production* of step  $k$ . Recall that  $D(\eta \| \zeta) := \text{tr}(\eta(\log \eta - \log \zeta))$  is the relative entropy, which satisfies  $D(\eta \| \zeta) \geq 0$  for  $\eta, \zeta \in \mathcal{D}(\mathcal{H})$ , and that  $S(\eta) := -\text{tr}[\eta \log \eta] \geq 0$  for  $\eta \in \mathcal{D}(\mathcal{H})$  is the von Neumann entropy. For notational simplicity, we at times omit the “i” superscript in  $\xi_k^i$ . Also, we omit tensored identities for operators acting trivially on the environment or on the system, whenever the context is clear.

These quantities are related through the *entropy balance equation*

$$\Delta S_k + \sigma_k = \beta_k \Delta Q_k, \quad (12.8)$$

(see e.g. [RW14; Han16b] for its brief derivation). This equation, together with  $\sigma_k \geq 0$ , i.e. the nonnegativity of the entropy production term, encapsulates the more general *Landauer principle*: when a system undergoes a state transformation by interacting with a thermal bath, the average increase in energy of the bath is bounded below by  $\beta^{-1}$  times the average decrease in entropy of the system. This principle was first presented in 1961 by Landauer [Lan61] and its saturation in quantum systems has more recently been investigated by Reeb and Wolf [RW14] and Jakšić and Pillet [JP14], the latter providing a treatment of the case of infinitely extended quantum systems.

If we consider a RIS with  $T$  steps, then summing (12.8) over  $k = 1, \dots, T$  yields the *total entropy balance equation*

$$\Delta S_{\mathcal{S},T} + \sigma_T^{\text{tot}} = \sum_{k=1}^T \beta_k \Delta Q_k, \quad (12.9)$$

where  $\Delta S_{\mathcal{S},T} = S(\rho^i) - S(\rho^f)$  and  $\rho^f = \rho_T$  is the state of  $\mathcal{S}$  after the final step of the RIS process (see (12.2)) and

$$\sigma_T^{\text{tot}} := \sum_{k=1}^T \sigma_k, \quad (12.10)$$

is the expected *total entropy production*.

## 12.4 The balance equation at the level of trajectories

We turn to obtaining an analogue of (12.9) for random variables on the probability space  $(\Omega_T, \mathcal{P}(\Omega_T), \mathbb{P}_T^F)$ . Remark first that  $\mathbb{P}_T^F(a^i, a^f, \vec{i}, \vec{j})$  and  $\mathbb{P}_T^B(a^i, a^f, \vec{i}, \vec{j})$  are of the form  $\mathbb{P}_T^F(a^i, a^f, \vec{i}, \vec{j}) = \text{tr}((\rho^i \otimes \Xi) S^* S)$  and  $\mathbb{P}_T^B(a^i, a^f, \vec{i}, \vec{j}) = \text{tr}((\rho_T^f \otimes \Xi) S S^*)$ . Under the assumption that  $\rho^i$  and  $\rho_T^f$  are faithful (full-rank) we therefore have

$$\mathbb{P}_T^F(a^i, a^f, \vec{i}, \vec{j}) = 0 \text{ if and only if } \mathbb{P}_T^B(a^i, a^f, \vec{i}, \vec{j}) = 0.$$

Since Proposition 10.3.7 shows that the image of a faithful state by an irreducible CPTP map is faithful,  $\rho^i$  and  $\rho_T^f$  will be faithful as long as  $\rho^i$  is faithful and each map  $\mathcal{L}_k$ , for  $k = 1, \dots, T$ , is irreducible.

This allows us to give the following definition:

**Definition 12.4.1.** If  $\rho^i$  and  $\rho_T^f$  are faithful, we define the classical random variable

$$\varsigma_T(a^i, a^f, \vec{i}, \vec{j}) := \log \frac{\mathbb{P}_T^F(a^i, a^f, \vec{i}, \vec{j})}{\mathbb{P}_T^B(a^i, a^f, \vec{i}, \vec{j})},$$

on  $(\Omega_T, \mathcal{P}(\Omega_T), \mathbb{P}_T^F)$ , which we call the entropy production of the repeated interaction system associated to the trajectory  $\omega = (a^i, a^f, \vec{i}, \vec{j})$ .

Note that the random variable  $\varsigma_T$  is the logarithm of the ratio of likelihoods, also known as the relative information random variable between  $\mathbb{P}_T^F$  and  $\mathbb{P}_T^B$  (see e.g. [CT06]). It is well-known that the distribution of such a random variable is related to the distinguishability of the two distributions (here  $\mathbb{P}_T^F$  and  $\mathbb{P}_T^B$ ): see e.g. [BD15]. Distinguishing between  $\mathbb{P}_T^F$  and  $\mathbb{P}_T^B$  amounts to testing the arrow of time; we refer the reader to [JOPS12; BJPP18] for a further discussion of this idea.

We have the following result, essentially present in [HP13].

**Lemma 12.4.2.** Assume  $\rho^i$  and  $\rho_T^f$  are faithful. If

$$i. \pi_{a^i}^i \rho^i \pi_{a^i}^i = \frac{\text{tr}(\rho^i \pi_{a^i}^i)}{\dim \pi_{a^i}^i} \pi_{a^i}^i \text{ for each } a^i,$$

$$ii. \pi_{a^f}^f \rho_T^f \pi_{a^f}^f = \frac{\text{tr}(\rho_T^f \pi_{a^f}^f)}{\dim \pi_{a^f}^f} \pi_{a^f}^f \text{ for each } a^f,$$

iii. for each  $k = 1, \dots, T$ , the state  $\xi_k$  (or equivalently  $h_{\mathcal{E}_k}$ ) is a function of  $Y_k$ ,

then

$$\varsigma_T(a^i, a^f, \vec{i}, \vec{j}) = \log \left( \frac{\text{tr}(\pi_{a^i}^i \rho^i)}{\text{tr}(\pi_{a^f}^f \rho_T^f)} \frac{\dim \pi_{a^f}^f}{\dim \pi_{a^i}^i} \right) + \sum_{k=1}^T \beta_k(E_{j_k}^{(k)} - E_{i_k}^{(k)}), \quad (12.11)$$

where  $E_{i_k}^{(k)} = \frac{\text{tr}(h\varepsilon_k \Pi_{i_k}^{(k)})}{\dim \Pi_{i_k}^{(k)}}$  are the energy levels of the  $k$ th probe.

*Remark.* The first two hypotheses are automatically satisfied if, for example,  $A^i$  and  $A^f$  are non-degenerate (all their spectral projectors are rank-one). All three hypotheses are automatically satisfied if, for example,  $\rho^i$ ,  $\rho_T^f$  and  $\xi_k$  can be written as functions of  $A^i$ ,  $A^f$  and  $Y_k$  (for each  $k = 1, \dots, T$ ) respectively.

*Proof of Lemma 12.4.2.* By definition

$$\begin{aligned} \varsigma_T(a^i, a^f, \vec{i}, \vec{j}) \\ = \log \frac{\text{tr}(U_T \cdots U_1(\pi_{a^i}^i \otimes \Pi_{\vec{i}})(\rho^i \otimes \Xi)(\pi_{a^i}^i \otimes \Pi_{\vec{i}})U_1^* \cdots U_T^*(\pi_{a^f}^f \otimes \Pi_{\vec{j}}))}{\text{tr}(U_1^* \cdots U_T^*(\pi_{a^f}^f \otimes \Pi_{\vec{j}})(\rho_T^f \otimes \Xi)(\pi_{a^f}^f \otimes \Pi_{\vec{j}})U_T \cdots U_1(\pi_{a^i}^i \otimes \Pi_{\vec{i}}))} \end{aligned}$$

using assumptions i., ii. and iii.,

$$\begin{aligned} &= \log \frac{\frac{\text{tr}(\rho^i \pi_{a^i}^i)}{\dim \pi_{a^i}^i} \prod_{k=1}^T \frac{\text{tr}(\xi_k \Pi_{i_k}^{(k)})}{\dim \Pi_{i_k}^{(k)}} \text{tr}(U_T \cdots U_1(\pi_{a^i}^i \otimes \Pi_{\vec{i}})U_1^* \cdots U_T^*(\pi_{a^f}^f \otimes \Pi_{\vec{j}}))}{\frac{\text{tr}(\rho_T^f \pi_{a^f}^f)}{\dim \pi_{a^f}^f} \prod_{k=1}^T \frac{\text{tr}(\xi_k \Pi_{j_k}^{(k)})}{\dim \Pi_{j_k}^{(k)}} \text{tr}(U_1^* \cdots U_T^*(\pi_{a^f}^f \otimes \Pi_{\vec{j}})U_T \cdots U_1(\pi_{a^i}^i \otimes \Pi_{\vec{i}}))} \\ &= \log \frac{\text{tr}(\rho^i \pi_{a^i}^i) \dim \pi_{a^f}^f}{\text{tr}(\rho_T^f \pi_{a^f}^f) \dim \pi_{a^i}^i} + \log \frac{\text{tr}(U_T \cdots U_1(\pi_{a^i}^i \otimes \Pi_{\vec{i}})U_1^* \cdots U_T^*(\pi_{a^f}^f \otimes \Pi_{\vec{j}}))}{\text{tr}(U_1^* \cdots U_T^*(\pi_{a^f}^f \otimes \Pi_{\vec{j}})U_T \cdots U_1(\pi_{a^i}^i \otimes \Pi_{\vec{i}}))} \\ &\quad + \log \prod_{k=1}^T \frac{\text{tr}(e^{-\beta_k h \varepsilon_k \Pi_{i_k}^{(k)}})}{Z_k \dim \Pi_{i_k}^{(k)}} - \log \prod_{k=1}^T \frac{\text{tr}(e^{-\beta_k h \varepsilon_k \Pi_{j_k}^{(k)}})}{Z_k \dim \Pi_{j_k}^{(k)}} \end{aligned}$$

and because  $\frac{\text{tr}(e^{-\beta_k h \varepsilon_k \Pi_{i_k}^{(k)}})}{\dim \Pi_{i_k}^{(k)}} = \exp\left(-\beta_k \frac{\text{tr}(e^{-\beta_k h \varepsilon_k \Pi_{i_k}^{(k)}})}{\dim \Pi_{i_k}^{(k)}}\right)$  by assumption iii. again,

$$\begin{aligned} &= \log \frac{\text{tr}(\rho^i \pi_{a^i}^i) \dim \pi_{a^f}^f}{\text{tr}(\rho_T^f \pi_{a^f}^f) \dim \pi_{a^i}^i} + \log \frac{\text{tr}(U_T \cdots U_1(\pi_{a^i}^i \otimes \Pi_{\vec{i}})U_1^* \cdots U_T^*(\pi_{a^f}^f \otimes \Pi_{\vec{j}}))}{\text{tr}(U_1^* \cdots U_T^*(\pi_{a^f}^f \otimes \Pi_{\vec{j}})U_T \cdots U_1(\pi_{a^i}^i \otimes \Pi_{\vec{i}}))} \\ &\quad + \sum_{k=1}^T \frac{\text{tr}(-\beta_k h \varepsilon_k \Pi_{i_k}^{(k)})}{\dim \Pi_{i_k}^{(k)}} - \sum_{k=1}^T \frac{\text{tr}(-\beta_k h \varepsilon_k \Pi_{j_k}^{(k)})}{\dim \Pi_{j_k}^{(k)}} \end{aligned}$$

and the last term vanishes by cyclicity of the trace.  $\square$

Again,  $\varsigma_T$  depends on the spectral projectors of the observables  $A^i$ ,  $A^f$  and  $(Y_k)_{k=1}^T$ , but not on their eigenvalues. However, with the choices  $A^i = -\log \rho^i$ ,  $A^f = -\log \rho_T^f$  and  $Y(s) = \beta(s)h\varepsilon(s)$ , and writing the spectral decompositions  $\rho^i = \sum r_a^i \pi_a^i$ ,  $\rho_T^f = \sum r_a^f \pi_a^f$ , the relation (12.11) takes the simpler form of a sum of differences of the obtained

eigenvalues (measurement results):

$$\varsigma_T(\omega) = (-\log r_{a^f}^f) - (-\log r_{a^i}^i) + \sum_{k=1}^T \beta_k (E_{j_k}^{(k)} - E_{i_k}^{(k)}),$$

which is the random variable introduced earlier as  $-\Delta a_T + \Delta y_T^{\text{tot}}(\omega)$  (again, in the case  $Y = \beta h_{\mathcal{E}}$ ). In this case,  $\Delta a_T = (-\log r_{a^i}^i) - (-\log r_{a^f}^f)$  is a classical random variable that is the difference of outcomes of measurements of entropy observables on the system  $\mathcal{S}$ , which we call  $\Delta s_{\mathcal{S},T}(\omega)$ . On the other hand,  $\sum_{k=1}^T \beta_k (E_{j_k}^{(k)} - E_{i_k}^{(k)})$  is a classical random variable that encapsulates Clausius' notion of the entropy increase of the chain  $(\mathcal{E}_k)_{k=1}^T$  at the level of trajectories, which we call  $\Delta s_{\mathcal{E},T}(\omega)$ . Then,

$$\Delta s_{\mathcal{S},T}(\omega) + \varsigma_T(\omega) = \Delta s_{\mathcal{E},T}(\omega), \quad (12.12)$$

and  $\varsigma_T(\omega)$  measures the difference between these two entropy variations, on the trajectory  $\omega$ .

Moreover, Proposition 12.4.3 below, whose proof is also left for the Appendix, links expression (12.11) to the entropy balance equation (12.9). Indeed, by showing that under suitable hypotheses the two terms on the right hand side of (12.11) average to the corresponding terms in (12.9), we show that  $\mathbb{E}_T(\varsigma_T) = \sigma_T^{\text{tot}}$ . In other words,  $\sigma_T^{\text{tot}}$  coincides with the (classical) relative entropy or Kullback–Leibler divergence  $D(\mathbb{P}_T^F \| \mathbb{P}_T^B)$  between the classical distributions  $\mathbb{P}_T^F$  and  $\mathbb{P}_T^B$ . Recall that  $D(\mathbb{P}_T^F \| \mathbb{P}_T^B) = 0$  if and only if  $\mathbb{P}_T^F = \mathbb{P}_T^B$ . Hence, we will refer to (12.11) as the *entropy balance equation at the level of trajectories*.

**Proposition 12.4.3.** *Assume that  $\rho^i$  is faithful and a function of  $A^i$ , that  $\rho_T^f$  is faithful and a function of  $A^f$ , and the state  $\xi_k$  (or equivalently  $h_{\mathcal{E}_k}$ ) is a function of  $Y_k$  for each  $k = 1, \dots, T$ , then*

$$\mathbb{E}_T \left[ \log \left( \frac{\text{tr}(\pi_{a^i}^i \rho^i) \dim \pi_{a^f}^f}{\text{tr}(\pi_{a^f}^f \rho^f) \dim \pi_{a^i}^i} \right) \right] = -\mathbb{E}_T(\Delta s_{\mathcal{S},T}) = S(\rho^f) - S(\rho^i) \quad (12.13)$$

and

$$\mathbb{E}_T \left( \sum_{k=1}^T \beta_k (E_{j_k}^{(k)} - E_{i_k}^{(k)}) \right) = \mathbb{E}_T(\Delta s_{\mathcal{E},T}) = \sum_{k=1}^T \beta_k \Delta Q_k. \quad (12.14)$$

Therefore,

$$\mathbb{E}_T(\varsigma_T) = \sigma_T^{\text{tot}}, \quad (12.15)$$

and relation (12.11) reduces to the entropy balance equation (12.9) upon taking expectation with respect to  $\mathbb{P}_T^F$ .

*Proof.* We start with the relation (12.13). On one hand,

$$\begin{aligned} \mathbb{E}_T \left( \log \frac{\text{tr}(\pi_{a^i}^i \rho^i)}{\dim \pi_{a^i}^i} \right) &= \sum_{a^i} \log \frac{\text{tr}(\pi_{a^i}^i \rho^i)}{\dim \pi_{a^i}^i} \sum_{a^f, \vec{i}, \vec{j}} \mathbb{P}_T^F(a^i, a^f, \vec{i}, \vec{j}) \\ &= \sum_{a^i} \log \frac{\text{tr}(\pi_{a^i}^i \rho^i)}{\dim \pi_{a^i}^i} \sum_{a^f, \vec{i}, \vec{j}} \text{tr} (U_T \cdots U_1 (\pi_{a^i}^i \rho^i \pi_{a^i}^i \otimes \Pi_{\vec{i}} \Xi \Pi_{\vec{i}}) U_1^* \cdots U_T^* (\pi_{a^f}^f \otimes \Pi_{\vec{j}})) \\ &= \sum_{a^i} \log \frac{\text{tr}(\pi_{a^i}^i \rho^i)}{\dim \pi_{a^i}^i} \sum_{\vec{i}} \text{tr} (U_T \cdots U_1 (\pi_{a^i}^i \rho^i \pi_{a^i}^i \otimes \Pi_{\vec{i}} \Xi \Pi_{\vec{i}}) U_1^* \cdots U_T^* (\mathbf{1} \otimes \mathbf{1})). \end{aligned}$$

Using that each  $\xi_k$  is a function of  $Y_k$ , we have  $\sum_{\vec{i}} \Pi_{\vec{i}} \Xi \Pi_{\vec{i}} = \Xi$ , and

$$\mathbb{E}_T \left( \log \frac{\text{tr}(\pi_{a^i}^i \rho^i)}{\dim \pi_{a^i}^i} \right) = \sum_{a^i} \log \frac{\text{tr}(\pi_{a^i}^i \rho^i)}{\dim \pi_{a^i}^i} \text{tr} (U_T \cdots U_1 (\pi_{a^i}^i \rho^i \pi_{a^i}^i \otimes \Xi) U_1^* \cdots U_T^*).$$

As  $\rho^i$  is a function of  $A^i$ , and using the cyclicity of the trace, we have

$$\begin{aligned} \mathbb{E}_T \left( \log \frac{\text{tr}(\pi_{a^i}^i \rho^i)}{\dim \pi_{a^i}^i} \right) &= \sum_{a^i} \log \frac{\text{tr}(\pi_{a^i}^i \rho^i)}{\dim \pi_{a^i}^i} \text{tr} (\pi_{a^i}^i \otimes \Xi) \frac{\text{tr}(\pi_{a^i}^i \rho^i)}{\dim \pi_{a^i}^i}, \\ &= \sum_{a^i} \text{tr}(\pi_{a^i}^i \rho^i) \log \frac{\text{tr}(\pi_{a^i}^i \rho^i)}{\dim \pi_{a^i}^i}. \end{aligned}$$

Using the commutation relation  $[\pi_{a^i}^i, \rho^i] = 0$ , the right-hand side is precisely  $-S(\rho^i)$ . The term involving  $\rho^f$  is treated similarly.

We turn to (12.14):

$$\begin{aligned} \mathbb{E}_T \left( \sum_{k=1}^T \beta_k (E_{j_k}^{(k)} - E_{i_k}^{(k)}) \right) &= \sum_{\vec{i}, \vec{j}} \sum_{k=1}^T \beta_k (E_{j_k}^{(k)} - E_{i_k}^{(k)}) \sum_{a^i, a^f} \mathbb{P}_T^F(a^i, a^f, \vec{i}, \vec{j}) \end{aligned}$$

using  $[\rho^i, A^i] = 0$ ,

$$\begin{aligned}
&= \sum_{\vec{i}, \vec{j}} \sum_{k=1}^T \beta_k (E_{j_k}^{(k)} - E_{i_k}^{(k)}) \operatorname{tr} (U_T \cdots U_1 (\rho^i \otimes \Pi_{\vec{i}} \Xi \Pi_{\vec{i}}) U_1^* \cdots U_T^* (\mathbf{1} \otimes \Pi_{\vec{j}})) \\
&= \sum_{\vec{j}} \sum_{k=1}^T \beta_k E_{j_k}^{(k)} \operatorname{tr} (U_T \cdots U_1 (\rho^i \otimes \Xi) U_1^* \cdots U_T^* (\mathbf{1} \otimes \Pi_{\vec{j}})) \\
&\quad - \sum_{\vec{i}} \sum_{k=1}^T \beta_k E_{i_k}^{(k)} \operatorname{tr} (U_T \cdots U_1 (\rho^i \otimes \Pi_{\vec{i}} \Xi \Pi_{\vec{i}}) U_1^* \cdots U_T^* (\mathbf{1} \otimes \mathbf{1})) \\
&= \sum_{k=1}^T \sum_{j_k} \beta_k \operatorname{tr} (U_T \cdots U_1 (\rho^i \otimes \Xi) U_1^* \cdots U_T^* (\mathbf{1} \otimes \Pi_{j_k}^{(k)} E_{j_k}^{(k)})) \\
&\quad - \sum_{k=1}^T \sum_{i_k} \beta_k \operatorname{tr} (\rho^i \otimes \Xi \Pi_{i_k}^{(k)} E_{i_k}^{(k)}).
\end{aligned}$$

Using that each  $\xi_k$  is a function of  $Y_k$ , we have  $\sum_{i_k} \Pi_{i_k}^{(k)} E_{i_k}^{(k)} = h_{\mathcal{E}_k}$  and thus

$$\mathbb{E}_T \left( \sum_{k=1}^T \beta_k (E_{j_k}^{(k)} - E_{i_k}^{(k)}) \right) = \sum_{k=1}^T \beta_k \operatorname{tr} (\xi_k^f h_{\mathcal{E}_k}) - \sum_{k=1}^T \beta_k \operatorname{tr} (\xi_k^i h_{\mathcal{E}_k}). \quad \square$$

Before we move on with our program, let us make a number of remarks on the choice of  $A^i = -\log \rho^i$  and  $A^f = -\log \rho_T^f$ .

*Remarks.*

- We made the assumption above that the operator  $A^f$  was uniformly bounded in  $T$ . The remarks below Corollary 12.7.2 show that this is true for  $A^f = -\log \rho_T^f$  in the case in which the reduced dynamics are identical ( $\mathcal{L} \equiv \mathcal{L}_k$ ) and irreducible, or in the case in which the  $\mathcal{L}_k$  are not all identical but are each irreducible and slowly vary in an adiabatic sense described below in Section 12.5.
- The observable  $-\log \rho^i$  is the analogue of the information random variable in classical information theory.
- The observable  $-\log \rho^f$  has the same interpretation and is the initial condition for the backward process in [EHM09; HP13] but it might seem odd that the observer is expected to have access to  $\rho^f = \mathcal{L}_T \circ \cdots \circ \mathcal{L}_1(\rho^i)$ . However, one can see that the reduced state of the probe after the forward experiment is random with  $\mathbb{P}_T^F$ -expectation equal to  $\rho^f$ . In addition,  $\varsigma_T$  is a relative information random



variable, and as such is relevant only to an observer who knows both distributions (here  $\mathbb{P}_T^F$  and  $\mathbb{P}_T^B$ ). Such an observer, knowing the possible outcomes for the random states after the experiment, and their distribution, would necessarily know their average,  $\rho^f$ .

We are interested in the full statistics of the random variable  $\varsigma_T$  that we will address through its cumulant generating functions in the limit  $T \rightarrow \infty$ . We will consider two cases:  $\lim_{T \rightarrow \infty} \sigma_T^{\text{tot}} < \infty$ , and  $\lim_{T \rightarrow \infty} \sigma_T^{\text{tot}} = \infty$ . The behaviour of this averaged quantity was investigated in [HJPR17] in an *adiabatic* setting in which the maps  $\mathcal{L}_k$  do not need to be identical, but instead are required to slowly vary (in a particular sense). This setting is described next.

## 12.5 Adiabatic limit of repeated interaction systems

[HJPR17] analyzed the Landauer principle and its saturation in the framework of an adiabatic limit of RIS that we briefly recall here. It is in this framework in which we will analyze  $\varsigma_T$ .

We introduce the *adiabatic parameter*  $T \in \mathbb{N}$  and consider a repeated interaction process with  $T$  probes, such that the parameters governing the  $k$ th probe and its interaction with  $\mathcal{S}$ , namely  $(h_{\mathcal{E}_k}, \beta_k, v_k)$ , are chosen by sampling sufficiently smooth functions, as described by the following assumption which we denote by the acronym **ADRS**, short for **adiabatic repeated interaction system**. Below, we say that a function  $f$  is  $C^2$  on  $[0, 1]$  if it is  $C^2$  on  $(0, 1)$ , and its first two derivatives admit limits at  $0^+$ ,  $1^-$ .

**ADRS** We are given a family of RIS processes indexed by an adiabatic parameter  $T \in \mathbb{N}$  such that there exist  $C^2$  functions  $s \mapsto h_{\mathcal{E}}(s)$ ,  $\beta(s)$ ,  $v(s)$  on  $[0, 1]$  for which

$$h_{\mathcal{E}_k} = h_{\mathcal{E}}\left(\frac{k}{T}\right), \quad \beta_k = \beta\left(\frac{k}{T}\right), \quad v_k = v\left(\frac{k}{T}\right)$$

for all  $k = 1, \dots, T$  when the adiabatic parameter has value  $T$ .

In this case, we may define

$$\begin{aligned} U(s) &= \exp\left(-i\tau\left(h_{\mathcal{E}}(s) + h_{\mathcal{E}}(s) + v(s)\right)\right), \\ \mathcal{L}(s) &= \text{tr}_{\mathcal{E}}\left(U(s)\left(\cdot \otimes \xi(s)\right)U(s)^*\right), \end{aligned} \tag{12.16}$$

where  $\xi(s)$  is the Gibbs state at inverse temperature  $\beta(s)$  for the Hamiltonian  $h_{\mathcal{E}}(s)$  and  $\tau$  is kept constant. Then,  $[0, 1] \ni s \mapsto \mathcal{L}(s)$  is a  $\mathcal{B}(\mathcal{B}(\mathcal{H}_{\mathcal{S}}))$ -valued  $C^2$  function, and

$\mathcal{L}_k = \mathcal{L}(\frac{k}{T})$  when the adiabatic parameter has value  $T$ . Note that for each  $s \in [0, 1]$ , the map  $\mathcal{L}(s)$  is completely positive (CP) and trace preserving (TP). For some results, we will need to make some extra hypotheses on the family  $(\mathcal{L}(s))_{s \in [0, 1]}$ . We introduce such conditions:

**Irr** For each  $s \in [0, 1]$ , the map  $\mathcal{L}(s)$  is *irreducible*, meaning that it has (up to a multiplicative constant) a unique invariant, which is a faithful state.

**Prim** For each  $s \in [0, 1]$ , the map  $\mathcal{L}(s)$  is *primitive*, meaning that it is irreducible and 1 is its only eigenvalue of modulus one.

From Proposition 10.3.8, we recall that the peripheral spectrum of an irreducible completely positive, trace-preserving map is a subgroup of the unit circle. We denote by  $z(s)$  the order of that subgroup for  $\mathcal{L}(s)$ .

The work [HJPR17] used a suitable adiabatic theorem to characterize the large  $T$  behaviour of the total entropy production term (12.10), which monitors the saturation of the Landauer bound in the adiabatic limit (note that the terms in the sum (12.10) are  $T$ -dependent through ADRIS). There, it was shown that for a RIS satisfying the assumptions ADRIS and Prim, the condition

$$\limsup_{T \rightarrow \infty} \sigma_T^{\text{tot}} < \infty$$

can be shown to be equivalent to the identity  $X(s) \equiv 0$ , where

$$X(s) := U(s)(\rho^{\text{inv}}(s) \otimes \xi^{\text{i}}(s))U(s)^* - \rho^{\text{inv}}(s) \otimes \xi^{\text{i}}(s), \quad (12.17)$$

and  $\rho^{\text{inv}}(s)$  is the unique invariant state of  $\mathcal{L}(s)$ . If the assumption  $X(s) \equiv 0$  does not hold, then  $\lim_{T \rightarrow \infty} \sigma_T^{\text{tot}} = \infty$ . The same work also established that the condition  $X(s) \equiv 0$  is equivalent to the existence of a family  $(k_{\mathcal{S}}(s))_{s \in [0, 1]}$  of observables on  $\mathcal{H}_{\mathcal{S}}$  such that  $[k_{\mathcal{S}}(s) + h_{\mathcal{E}}(s), U(s)] \equiv 0$ .

Let us briefly summarize the proof from [HJPR17], in high level terms. The goal is to estimate  $\sigma_T^{\text{tot}} = \sum_{k=1}^T \sigma_k$ . From (12.7) we have

$$\sigma_k \equiv \sigma_{k,T} = D(U(\frac{k}{T})(\rho_{k-1} \otimes \xi(\frac{k}{T}))U(\frac{k}{T})^* \| \mathcal{L}(\frac{k}{T})(\rho_{k-1}) \otimes \xi(\frac{k}{T})) \quad (12.18)$$

with

$$\rho_{k-1} = \mathcal{L}(\frac{k-1}{T}) \circ \dots \circ \mathcal{L}(\frac{1}{T})(\rho^{\text{i}}). \quad (12.19)$$

Note then that in  $\sigma_k$ , each quantity depends only on the current step, i.e., on the various quantities defining the repeated interaction system evaluated at  $s = \frac{k}{T}$ , except for  $\rho_{k-1}$ ,

which depends on all the previous timesteps  $\{\frac{1}{T}, \frac{2}{T}, \dots, \frac{k-1}{T}\}$  as well. The proof therefore proceeds by estimating  $\rho_{k-1} = \rho^{\text{inv}}(\frac{k-1}{T}) + O(1/T) = \rho^{\text{inv}}(\frac{k}{T}) + O(1/T)$  where  $\rho^{\text{inv}}(s)$  is the unique invariant state of  $\mathcal{L}(s)$ , using a discrete adiabatic theorem, to remove the dependence on the previous timesteps. This is the main place where the assumptions [ADRIS](#) and [Prim](#) are used in the proof. Then, roughly speaking, a perturbation theory argument is used to control the relative entropy term  $\sigma_k$  in terms of  $\|X(k/T)\|$ , yielding in the end that if  $\|X(s)\| \equiv 0$ ,  $\sigma^{\text{tot}}$  vanishes (under an assumption on the initial state), and otherwise diverges.

In this work, we aim to refine this analysis by investigating the entropy production at the level of trajectories, i.e. study the random variable  $\varsigma_T$  (which satisfies  $\mathbb{E}[\varsigma_T] = \sigma_T^{\text{tot}}$  by Proposition [12.4.3](#)). It turns out that by considering the moment-generating function of  $\varsigma_T$ , a similar proof technique to the one sketched above for  $\sigma_T^{\text{tot}}$  can be used.

## 12.6 Moment generating functions and deformed CP maps

First, we add an additional assumption about the observable  $Y$  which is measured as part of the two-time measurement protocol. When assuming [ADRIS](#) and discussing measured observables  $Y$ , we will always assume the following:

**Comm** There is a twice continuously differentiable  $\mathcal{B}(\mathcal{H}_{\mathcal{E}})$ -valued function  $s \mapsto Y(s)$  on  $[0, 1]$  such that  $[Y(s), h_{\mathcal{E}}(s)] = 0$  at all  $s \in [0, 1]$  for which, when the adiabatic parameter has value  $T$ ,

$$Y_k = Y(\frac{k}{T}), \quad k = 1, \dots, T.$$

The family of probe Hamiltonians themselves  $Y(s) = h_{\mathcal{E}}(s)$  are suitable, but in our applications to Landauer's Principle, we will be particularly interested in  $Y(s) = \beta(s)h_{\mathcal{E}}(s)$ .

We recall that the quantities  $\Delta a_T$  and  $\Delta y_T^{\text{tot}}$  are defined in [\(12.5\)](#) and [\(12.6\)](#). We also recall that the *moment generating function* of a real-valued random variable  $V$  (with respect to the probability distribution  $\mathbb{P}_T^F$ , which will always be implicit in this chapter) is defined as the map  $M_V : \alpha \mapsto \mathbb{E}_T(e^{\alpha V})$ , and the moment generating function of a pair  $(V_1, V_2)$  as the map  $M_{(V_1, V_2)} : (\alpha_1, \alpha_2) \mapsto \mathbb{E}_T(e^{\alpha_1 V_1 + \alpha_2 V_2})$ . When  $V$  or  $(V_1, V_2)$  are given by the random variables  $\Delta y_T^{\text{tot}}$ ,  $\Delta a_T$ , the above functions  $M_V$  (resp.  $M_{(V_1, V_2)}$ ) are defined for all  $\alpha \in \mathbb{C}$  (resp. for all  $(\alpha_1, \alpha_2) \in \mathbb{C}^2$ ). For relevant properties of moment generating functions we refer the reader to Sections 21 and 30 of [\[Bil95\]](#).

Our main tool to study these moment generating functions is the following proposition:

**Proposition 12.6.1.** *For  $\alpha \in \mathbb{C}$ , define an analytic deformation of  $\mathcal{L}(s)$  by the complex parameter  $\alpha$  corresponding to the observable  $Y(s)$ :*

$$\begin{aligned} \mathcal{L}_Y^{(\alpha)}(s) : \mathcal{B}(\mathcal{H}_S) &\rightarrow \mathcal{B}(\mathcal{H}_S) \\ \eta &\mapsto \text{tr}_{\mathcal{E}} \left( e^{\alpha Y(s)} U(s) (\eta \otimes \xi(s)) e^{-\alpha Y(s)} U(s)^* \right). \end{aligned} \quad (12.20)$$

Under assumption [Comm](#), the moment generating function of  $\Delta y_T^{\text{tot}}$  is given by

$$M_{\Delta y_T^{\text{tot}}}(\alpha) = \text{tr}_S \left( \mathcal{L}_Y^{(\alpha)}\left(\frac{T}{T}\right) \cdots \mathcal{L}_Y^{(\alpha)}\left(\frac{1}{T}\right) \left( \sum_{a^i} \pi_{a^i}^i \rho^i \pi_{a^i}^i \right) \right).$$

If in addition  $[A^i, \rho^i] = 0$ , then the moment generating function of the pair  $(\Delta y_T^{\text{tot}}, \Delta a_T)$  is given by

$$M_{(\Delta y_T^{\text{tot}}, \Delta a_T)}(\alpha_1, \alpha_2) = \text{tr} \left( e^{-\alpha_2 A^f} \mathcal{L}_Y^{(\alpha_1)}\left(\frac{T}{T}\right) \cdots \mathcal{L}_Y^{(\alpha_1)}\left(\frac{1}{T}\right) (e^{+\alpha_2 A^i} \rho^i) \right).$$

so that in particular the moment generating function of  $-\Delta a_T + \Delta y_T^{\text{tot}}$  is given by

$$M_{-\Delta a_T + \Delta y_T^{\text{tot}}}(\alpha) = \text{tr}_S \left( e^{+\alpha A^f} \mathcal{L}_Y^{(\alpha)}\left(\frac{T}{T}\right) \cdots \mathcal{L}_Y^{(\alpha)}\left(\frac{1}{T}\right) (e^{-\alpha A^i} \rho^i) \right). \quad (12.21)$$

*Remarks.*

- The complex deformation of the map  $\mathcal{L}(s)$  given here is similar to the deformations introduced in [\[HMO07\]](#) for hypothesis testing on spin chains, and to the complex deformation of Lindblad operators introduced in [\[JPW14\]](#) suited to the study of entropy fluctuations for continuous time evolution.
- With the choices  $Y(s) = \beta(s)h_{\mathcal{E}}(s)$ ,  $A^i = -\log \rho^i$ , and  $A^f = -\log \rho_T^f$ , the balance equation [\(12.12\)](#) yields  $\varsigma_T = -\Delta a_T + \Delta y_T^{\text{tot}}$ . Thus, in this case [\(12.21\)](#) provides a formula for the moment generating function of  $\varsigma_T$  in terms of the composition  $\mathcal{L}_Y^{(\alpha)}\left(\frac{T}{T}\right) \cdots \mathcal{L}_Y^{(\alpha)}\left(\frac{1}{T}\right) (e^{-\alpha A^i} \rho^i)$ . In particular, the dependence of  $\varsigma_T$  on the different timesteps of the evolution is captured entirely in that composition, just as for the averaged quantity  $\sigma_k$  in [\(12.18\)](#). This is the key step that allows an adiabatic theorem to approximate this composition in order to analyze  $\varsigma_T$  in the limit  $T \rightarrow \infty$ .

*Proof.* [Comm](#) yields that  $[Y(s), \xi(s)] = 0$  for all  $s$ . Then by the expression [\(12.3\)](#),

$$\mathbb{E}(e^{\alpha \Delta y_T^{\text{tot}}})$$

$$\begin{aligned}
&= \sum_{\vec{i}, \vec{j}} \sum_{a^i, a^f} \exp \left( \alpha \sum_{k=1}^Y (y_{j_k}^{(k)} - y_{i_k}^{(k)}) \right) \mathbb{P}_T^F(a^i, a^f, \vec{i}, \vec{j}) \\
&= \sum_{\vec{i}, \vec{j}} \text{tr} \left( U_T \cdots U_1 \left( \sum_{a^i} \pi_{a^i}^i \rho^i \pi_{a^i}^i \otimes \prod_{k=1}^T e^{-\alpha y_{i_k}^{(k)}} \Pi_{i_k} \Xi \right) U_1^* \cdots U_T^* \left( \sum_b \pi_{a^f}^f \otimes \prod_{k=1}^T e^{\alpha y_{j_k}^{(k)}} \Pi_{j_k} \right) \right) \\
&= \text{tr} \left( U_T \cdots U_1 \left( \sum_{a^i} \pi_{a^i}^i \rho^i \pi_{a^i}^i \otimes \prod_{k=1}^T e^{-\alpha Y_k} \Xi \right) U_1^* \cdots U_T^* \left( \text{id} \otimes \prod_{k=1}^T e^{\alpha Y_k} \right) \right) \\
&= \text{tr} \left( \mathcal{L}_Y^{(\alpha)} \left( \frac{T}{T} \right) \circ \cdots \circ \mathcal{L}_Y^{(\alpha)} \left( \frac{1}{T} \right) \left( \sum_{a^i} \pi_{a^i}^i \rho^i \pi_{a^i}^i \right) \right).
\end{aligned}$$

Assume in addition that  $[A^i, \rho^i] = 0$ . Then similarly, for  $\alpha_1, \alpha_2$  in  $\mathbb{R}$ ,

$$\begin{aligned}
&\mathbb{E}(e^{\alpha_1 \Delta y_T^{\text{tot}} + \alpha_2 \Delta a_T}) \\
&= \sum_{\vec{i}, \vec{j}} \sum_{a^i, a^f} \exp \left( \alpha_1 \sum_{k=1}^Y (y_{j_k}^{(k)} - y_{i_k}^{(k)}) \right) e^{\alpha_2 (a^i - a^f)} \mathbb{P}_T^F(a^i, a^f, \vec{i}, \vec{j}) \\
&= \sum_{\vec{i}, \vec{j}} \text{tr} \left( U_T \cdots U_1 \left( \sum_{a^i} e^{+\alpha_2 a^i} \pi_{a^i}^i \rho^i \otimes \prod_{k=1}^T e^{-\alpha_1 y_{i_k}^{(k)}} \Pi_{i_k} \Xi \right) \right. \\
&\quad \left. U_1^* \cdots U_T^* \left( \sum_b e^{-\alpha_2 a^f} \pi_{a^f}^f \otimes \prod_{k=1}^T e^{\alpha_1 y_{j_k}^{(k)}} \Pi_{j_k} \right) \right) \\
&= \text{tr} \left( U_T \cdots U_1 (e^{+\alpha_2 A^i} \rho^i \otimes \prod_{k=1}^T e^{-\alpha_1 Y_k} \Xi) U_1^* \cdots U_T^* (e^{-\alpha_2 A^f} \otimes \prod_{k=1}^T e^{\alpha_1 Y_k}) \right) \\
&= \text{tr} \left( e^{-\alpha_2 A^f} \mathcal{L}_Y^{(\alpha_1)} \left( \frac{T}{T} \right) \circ \cdots \circ \mathcal{L}_Y^{(\alpha_1)} \left( \frac{1}{T} \right) (e^{+\alpha_2 A^i} \rho^i) \right). \quad \square
\end{aligned}$$

Section 3.2 [HJPR18] establishes several properties of the maps  $\mathcal{L}_Y^{(\alpha)}(s)$  for  $s \in [0, 1]$  and  $\alpha \in \mathbb{C}$ , including that they are each completely positive and irreducible (under the assumption Irr), although in general not trace-preserving. Appendix A of the same work goes on to analyze the decomposition of irreducible maps given in Proposition 10.3.8 in the presence of the deformation parameter  $\alpha \in \mathbb{R}$ . The following summarizes the results of that work.

**Proposition 12.6.2.** *Let  $s \mapsto \mathcal{L}(s)$  be a family of CPTP maps satisfying Irr, and define  $\mathcal{L}_Y^{(\alpha)}(s)$  by (12.20). Then there exist maps  $\lambda_Y^{(\alpha)}(s)$ ,  $I_Y^{(\alpha)}(s)$  and  $\rho_Y^{(\alpha)}(s)$  from  $[0, 1] \times \mathbb{R} \ni (s, \alpha)$  to, respectively,  $\mathbb{R}_{>0}$ , the set of positive-definite operators, and the set of faithful states of  $\mathcal{H}_S$ ; and maps  $z(s), u(s)$  from  $[0, 1]$  to, respectively,  $\mathbb{N}$  and the set of unitary operators, with the following properties:*

- the identities  $[u(s), I_Y^{(\alpha)}(s)] = [u(s), \rho_Y^{(\alpha)}(s)] = 0$ , and  $u(s)^{z(s)} = \text{id}$  hold;
- the peripheral spectrum of  $\mathcal{L}_Y^{(\alpha)}(s)$  is  $\lambda_Y^{(\alpha)}(s)S_{z(s)}$ , where

$$S_z = \{\theta^m : \theta = e^{2i\pi/z}, m = 0, \dots, z-1\};$$

- the spectral decomposition  $u(s) = \sum_{m=1}^{z(s)} e^{2i\pi m/z(s)} p_m(s)$  holds;
- the map  $\eta \mapsto \text{tr}(I_Y^{(\alpha)}(s)u(s)^{-m}\eta)\rho_Y^{(\alpha)}(s)u(s)^m$  is the spectral projector of  $\mathcal{L}_Y^{(\alpha)}(s)$  associated with  $\lambda_Y^{(\alpha)}(s)e^{2i\pi m/z(s)}$ ;
- the unitary  $u(s)$  and cardinal  $z(s)$  of the peripheral spectrum of  $\mathcal{L}_Y^{(\alpha)}(s)$  do not depend on  $\alpha$  or  $Y$ .

Note that since  $\mathcal{L}^{(0)}(s) = \mathcal{L}(s)$ , we have  $\lambda_Y^{(0)}(s) = 1$ ,  $I_Y^{(0)}(s) = \text{id}$  and  $\rho_Y^{(0)}(s) = \rho^{\text{inv}}(s)$  for all  $Y$  and  $s$ . As mentioned above, the case  $Y = \beta h_{\mathcal{E}}$  will be particularly relevant to the discussion of the Landauer's Principle. We therefore drop the indices  $Y$ , and simply denote by  $\lambda^{(\alpha)}(s)$ ,  $I^{(\alpha)}(s)$  and  $\rho^{(\alpha)}(s)$  the above quantities, in the case where  $Y = \beta h_{\mathcal{E}}$ . We also define the rescaled map

$$\tilde{\mathcal{L}}_Y^{(\alpha)}(s) = (\lambda_Y^{(\alpha)}(s))^{-1} \mathcal{L}_Y^{(\alpha)}(s)$$

which has spectral radius 1.

## 12.7 An adiabatic theorem

The following proposition establishes an adiabatic theorem for compositions of the form  $\tilde{\mathcal{L}}_Y^{(\alpha)}(\frac{k}{T}) \dots \tilde{\mathcal{L}}_Y^{(\alpha)}(\frac{1}{T}) \rho^i$  in terms of the quantities given in Proposition 12.6.2. See Proposition 3.12 in [HJPR18] for the proof, which is based on the proof of the adiabatic theorem given in [HJPR17].

**Proposition 12.7.1.** *Consider an ADRIS with the family  $(\mathcal{L}(s))_{s \in [0,1]}$  satisfying Irr with  $z(s) \equiv z$ . Then, there exist continuous functions  $\mathbb{R} \ni \alpha \mapsto \ell'(\alpha) \in (0,1)$  and  $\mathbb{R} \ni \alpha \mapsto C(\alpha) \in \mathbb{R}_+$ , and a function  $\alpha \mapsto T_0(\alpha) \in \mathbb{N}$  that is bounded on any compact set of  $\mathbb{R}$ , such that for all  $\alpha \in \mathbb{R}$ ,  $T \geq T_0(\alpha)$ , and  $k \leq T$ ,*

$$\begin{aligned} \left\| \tilde{\mathcal{L}}_Y^{(\alpha)}(\frac{k}{T}) \dots \tilde{\mathcal{L}}_Y^{(\alpha)}(\frac{1}{T}) \rho^i - z e^{-\vartheta_Y^{(\alpha)}} \sum_{m=0}^{z-1} \text{tr}(I^{(\alpha)}(0) p_m(0) \rho^i) \rho_Y^{(\alpha)}(\frac{k}{T}) p_{m-k}(\frac{k}{T}) \right\| \\ \leq \frac{C(\alpha)}{T(1 - \ell'(\alpha))} + C(\alpha) \ell'(\alpha)^k \end{aligned}$$

where the index of the spectral projector  $p_{m-k}(\frac{k}{T})$  is interpreted modulo  $z$ , and

$$\vartheta_Y^{(\alpha)} := \int_0^{k/T} \text{tr} \left( I_Y^{(\alpha)}(s) \frac{\partial}{\partial s} \rho_Y^{(\alpha)}(s) \right) ds.$$

By taking  $\alpha = 0$ , this result allows adiabatic approximation of the state of  $\mathcal{S}$  under the physical evolution  $\mathcal{L}(\frac{k}{T}) \cdots \mathcal{L}(\frac{1}{T})$  after  $k$  steps of an irreducible RIS. This corresponds to a generalization of the results of [HJPR17], which could only treat the primitive case, i.e.  $z = 1$ .

**Corollary 12.7.2.** *Consider an ADRIS with the family  $(\mathcal{L}(s))_{s \in [0,1]}$  satisfying Irr with  $z(s) \equiv z$ . Then, there exists  $\ell' < 1$ ,  $C > 0$ , and  $T_0 > 0$  such that for all  $T \geq T_0$ , and  $k \leq T$ ,*

$$\left\| \mathcal{L}\left(\frac{k}{T}\right) \cdots \mathcal{L}\left(\frac{1}{T}\right) \rho^i - \rho_{\text{adiab}}(k, T) \right\| \leq \frac{C}{T(1 - \ell')} + C\ell'^k$$

where

$$\rho_{\text{adiab}}(k, T) := z \sum_{n=0}^{z-1} \text{tr} \left( p_n(0) \rho^i \right) \rho^{\text{inv}}\left(\frac{k}{T}\right) p_{n-k}\left(\frac{k}{T}\right) \quad (12.22)$$

is a state, and the index of the spectral projector  $p_{n-k}(\frac{k}{T})$  is interpreted modulo  $z$ . Moreover, if  $\rho^i$  is faithful, we have the uniform bound

$$\inf_{T > 1} \inf_{k \leq T} \inf \text{spec} \rho_{\text{adiab}}(k, T) \geq z \left( \min_{1 \leq j \leq z} \text{tr} \left( p_j(0) \rho^i \right) \right) \inf_{s \in [0,1]} \inf \text{spec} \rho^{\text{inv}}(s) > 0.$$

*Proof.* We apply Proposition 12.7.1 for  $\alpha = 0$ , and use that  $I^{(0)}(s) \equiv \text{id}$ ,  $\rho_Y^{(0)} = \rho^{\text{inv}}$ , and  $\vartheta_Y^{(\alpha)} = 0$  which follows from  $\text{tr} \rho_Y^{(0)}(s) \equiv 1$ . Next, we check the formula  $\text{tr}(\rho^{\text{inv}}(\frac{k}{T}) p_\ell(\frac{k}{T})) = \frac{1}{z}$  for each  $\ell = 0, \dots, z-1$ . We drop the argument  $\frac{k}{T}$  in what follows, and write  $\mathcal{L}_{k/T}(\cdot) = \sum_i V_i \cdot V_i^*$  the Kraus decomposition. Recalling that  $p_\ell V_i = V_i p_{\ell+1}$  for all  $i$  and  $\ell$  via (10.25),

$$\begin{aligned} \text{tr}(\rho^{\text{inv}} p_\ell) &= \text{tr}(\mathcal{L}(\rho^{\text{inv}}) p_\ell) = \sum_i \text{tr}(V_i \rho^{\text{inv}} V_i^* p_\ell) \\ &= \sum_i \text{tr}(V_i \rho^{\text{inv}} p_{\ell+1} V_i^*) = \text{tr}(\mathcal{L}(\rho^{\text{inv}} p_{\ell+1})), \end{aligned}$$

so  $\text{tr}(\rho^{\text{inv}} p_\ell) = \text{tr}(\rho^{\text{inv}} p_{\ell+1})$  using that  $\mathcal{L}$  is trace-preserving. As  $\sum_\ell \text{tr}(\rho^{\text{inv}} p_\ell) = \text{tr} \rho^{\text{inv}} = 1$ , we must have  $\text{tr}(\rho^{\text{inv}} p_\ell) = \frac{1}{z}$ . Therefore,

$$\text{tr}(\rho_{\text{adiab}}(k, T)) = z \sum_n \text{tr} \left( p_n(0) \rho^i \right) \text{tr} \left( \rho^{\text{inv}}\left(\frac{k}{T}\right) p_{n-k}\left(\frac{k}{T}\right) \right) = \sum_n \text{tr} \left( p_n(0) \rho^i \right)$$

$$= \text{tr } \rho^i = 1.$$

Moreover, given a normalized vector  $\psi \in \mathcal{H}$ , we have

$$\begin{aligned} \langle \psi, \rho_{\text{adiab}}(k, T) \psi \rangle &= z \sum_n \text{tr} (p_n(0) \rho^i) \langle \psi, \rho^{\text{inv}}(\frac{k}{T}) p_{n-k}(\frac{k}{T}) \psi \rangle \\ &= z \sum_n \text{tr} (p_n(0) \rho^i) \langle p_{n-k}(\frac{k}{T}) \psi, \rho^{\text{inv}}(\frac{k}{T}) p_{n-k}(\frac{k}{T}) \psi \rangle. \end{aligned}$$

using  $[\rho^{\text{inv}}(\frac{k}{T}), p_{n-k}(\frac{k}{T})] = 0$ . Since  $\rho^{\text{inv}}(\frac{k}{T}) > 0$  and  $\text{tr} (p_n(0) \rho^i) > 0$ , each term in the sum is non-negative, and we have

$$\begin{aligned} \langle \psi, \rho_{\text{adiab}}(k, T) \psi \rangle &\geq z \left( \min_{1 \leq j \leq z} \text{tr} (p_j(0) \rho^i) \right) \sum_n \langle \psi, \rho^{\text{inv}}(\frac{k}{T}) p_{n-k}(\frac{k}{T}) \psi \rangle \\ &= z \left( \min_{1 \leq j \leq z} \text{tr} (p_j(0) \rho^i) \right) \langle \psi, \rho^{\text{inv}}(\frac{k}{T}) \psi \rangle \\ &\geq z \left( \min_{1 \leq j \leq z} \text{tr} (p_j(0) \rho^i) \right) \inf_{s \in [0,1]} \inf \text{spec } \rho^{\text{inv}}(s). \quad \square \end{aligned}$$

*Remarks.*

- Given an [ADRS](#) the family  $(\mathcal{L}(s))_{s \in [0,1]}$  satisfying [Irr](#), for faithful  $\rho^i$  the state  $\rho_T^f = \mathcal{L}_T \cdots \mathcal{L}_1 \rho^i$  is faithful for each  $T > 1$  using Proposition [10.3.7](#). Corollary [12.7.2](#) and Weyl's inequalities (see Section III.2 in [\[Bha97\]](#)) give the stronger result  $\inf_{T>1} \inf \text{spec } \rho_T^f > 0$ . In particular, we may make the choice  $A^f = -\log \rho_T^f$  which is bounded uniformly in  $T$ .
- If we assume, in the notation of Proposition [10.3.8](#), that  $P_n \rho^i = 0$  for  $n > 0$  (i.e.  $\rho^i$  has no components corresponding to the peripheral eigenvalues of  $\mathcal{L}(0)$  other than 1), then one can check that  $\rho_{\text{adiab}}(k, T) = \rho^{\text{inv}}(\frac{k}{T})$ .

## 12.8 The statistics of the entropy production at the level of trajectories

In this section, we summarize the statistical properties of  $\zeta_T$  which are derived in [\[HJPR18\]](#) using the adiabatic theorem described in Section [12.7](#). For the sake of brevity, most proofs will be omitted.

First, the adiabatic theorem can be used to directly compute the limiting moment generating function in the special case in which the reduced dynamics is primitive and  $X(s) \equiv 0$ .



**Theorem 12.8.1.** *Under the assumptions [ADRI](#) and [Prim](#), and with  $X(s) \equiv 0$ , the limiting moment generating function  $M_\varsigma(\alpha) := \lim_{T \rightarrow \infty} M_{\varsigma_T}(\alpha)$  satisfies*

$$M_\varsigma(\alpha) = Q_{-\alpha}(\rho^{\text{inv}}(0) \parallel \rho^{\text{i}}),$$

for  $Q_\alpha(\eta \parallel \zeta) := \text{tr}(\eta^\alpha \zeta^{1-\alpha})$  which can be related to the Rényi relative entropy via  $D_\alpha(\eta \parallel \zeta) = \frac{1}{1-\alpha} \log Q_\alpha(\eta \parallel \zeta)$ .

*Remark.* In particular, if  $\rho^{\text{i}}$  is chosen as the initial invariant state  $\rho^{\text{inv}}(0)$ , then  $M_\varsigma(\alpha) \equiv 1$ , and hence  $\varsigma_T \rightarrow 0$  weakly. This greatly strengthens the results of [\[HJPR17\]](#) which showed that under the assumptions of the proposition, if  $\rho^{\text{i}} = \rho^{\text{inv}}(0)$  then  $\sigma_T^{\text{tot}} = \mathbb{E}[\varsigma_T] \rightarrow 0$ .

Next, we remove the assumptions of primitivity and that  $X(s) \equiv 0$ , and use the adiabatic theorem to compute the exponential rate of change of the moment generating function with  $T$  as  $T \rightarrow \infty$ .

**Lemma 12.8.2.** *Under the assumptions [ADRI](#) and [Irr](#) with  $z(s) \equiv z$ , for any faithful initial state  $\rho^{\text{i}} > 0$ , for any  $\alpha \in \mathbb{R}$ , the moment generating function of the random variable  $\Delta y_T^{\text{tot}}$  with respect to  $\mathbb{P}_T^F$  satisfies*

$$\lim_{T \rightarrow \infty} \frac{1}{T} \log M_{\Delta a_T + \Delta y_T^{\text{tot}}}(\alpha) = \lim_{T \rightarrow \infty} \frac{1}{T} \log M_{\Delta y_T^{\text{tot}}}(\alpha) = \int_0^1 \log \lambda_Y^{(\alpha)}(s) \, ds =: \Lambda_Y(\alpha). \quad (12.23)$$

*Proof.* Lemma 5.1 of [\[HJPR18\]](#) uses the adiabatic theorem Proposition [12.7.1](#) to show that

$$\lim_{T \rightarrow \infty} \frac{1}{T} \log \text{tr} \left( \tilde{\mathcal{L}}_Y^{(\alpha)}\left(\frac{T}{T}\right) \cdots \tilde{\mathcal{L}}_Y^{(\alpha)}\left(\frac{1}{T}\right) (\rho^{\text{i}}) \right) = 0.$$

But the moment generating function reads

$$\begin{aligned} M_{\Delta y_T^{\text{tot}}}(\alpha) &= \text{tr} \left( \mathcal{L}_Y^{(\alpha)}\left(\frac{T}{T}\right) \cdots \mathcal{L}_Y^{(\alpha)}\left(\frac{1}{T}\right) \rho^{\text{i}} \right) \\ &= \left( \prod_{k=1}^T \lambda_Y^{(\alpha)}\left(\frac{k}{T}\right) \right) \text{tr} \left( \tilde{\mathcal{L}}_Y^{(\alpha)}\left(\frac{T}{T}\right) \cdots \tilde{\mathcal{L}}_Y^{(\alpha)}\left(\frac{1}{T}\right) (\rho^{\text{i}}) \right) \end{aligned}$$

by definition of  $\tilde{\mathcal{L}}_Y^{(\alpha)}(s)$ . Hence, the result follows from the Riemann sum convergence

$$\begin{aligned} \lim_{T \rightarrow \infty} \frac{1}{T} \log \left( \prod_{k=1}^T \lambda_Y^{(\alpha)}\left(\frac{k}{T}\right) \right) &= \lim_{T \rightarrow \infty} \sum_{k=1}^T \left( \frac{k}{T} - \frac{k-1}{T} \right) \log \lambda_Y^{(\alpha)}\left(\frac{k}{T}\right) \\ &= \int_0^1 \log \lambda_Y^{(\alpha)}(s) \, ds. \end{aligned}$$

The same holds for  $M_{\Delta a_T + \Delta y_T^{\text{tot}}}(\alpha)$  because  $\Delta y_T^{\text{tot}}(\omega)$  and  $-\Delta a_T + \Delta y_T^{\text{tot}}(\omega)$  only differ by the uniformly bounded term  $-\Delta a_T(\omega) + \Delta y_T^{\text{tot}}(\omega)$ .  $\square$

Many results can be derived from the convergence of the moment generating function of  $\varsigma_T$  shown in Lemma 12.8.2. First, we formulate a large deviations principle; see e.g. [Ell06] for more on large deviations. In the statement below we denote the interior of a set  $E \subseteq \mathbb{R}$  by  $\text{int } E$  and the closure by  $\text{cl } E$ .

**Theorem 12.8.3.** *Assume ADRIS and Irr, and that the initial state  $\rho^i$  is faithful. Let  $\Lambda_Y$  be defined by relation (12.23) and denote by  $\Lambda_Y^*$  the Fenchel–Legendre transform of  $\Lambda_Y$ , i.e. for  $x \in \mathbb{R}$  let*

$$\Lambda_Y^*(x) = \sup_{\alpha \in \mathbb{R}} (\alpha x - \Lambda_Y(\alpha)).$$

*Then  $\Lambda_Y^*(x) = +\infty$  for  $x \notin [\nu_{Y,-}, \nu_{Y,+}]$ , and for any Borel set  $E$  of  $\mathbb{R}$  one has*

$$\begin{aligned} - \inf_{x \in \text{int } E} \Lambda_Y^*(x) &\leq \liminf_{T \rightarrow \infty} \frac{1}{T} \log \mathbb{P}_T^F \left( \frac{\Delta y_T^{\text{tot}}}{T} \in \text{int } E \right) \\ &\leq \limsup_{T \rightarrow \infty} \frac{1}{T} \log \mathbb{P}_T^F \left( \frac{\Delta y_T^{\text{tot}}}{T} \in \text{cl } E \right) \leq - \inf_{x \in \text{cl } E} \Lambda_Y^*(x). \end{aligned}$$

*The same statement holds with  $-\Delta a_T + \Delta y_T^{\text{tot}}$  in place of  $\Delta y_T^{\text{tot}}$ . In particular, for  $Y = \beta h_{\mathcal{E}}$ , one has*

$$\begin{aligned} - \inf_{x \in \text{int } E} \Lambda^*(x) &\leq \liminf_{T \rightarrow \infty} \frac{1}{T} \log \mathbb{P}_T^F \left( \frac{\varsigma_T}{T} \in \text{int } E \right) \\ &\leq \limsup_{T \rightarrow \infty} \frac{1}{T} \log \mathbb{P}_T^F \left( \frac{\varsigma_T}{T} \in \text{cl } E \right) \leq - \inf_{x \in \text{cl } E} \Lambda^*(x) \end{aligned}$$

*and the same statement holds with  $\Delta s_{\mathcal{E},T}$  in place of  $\varsigma_T$ .*

Let us explore two special cases of this result with the choice  $Y(s) = \beta(s)h_{\mathcal{E}}(s)$ . First, if  $X(s) \equiv 0$ , one can compute<sup>1</sup> that  $\lambda^{(\alpha)}(s) = 1$  for all  $\alpha \in \mathbb{R}$  and  $s \in [0, 1]$ . In that case  $\Lambda(\alpha) \equiv 0$  and

$$\Lambda^*(x) = \begin{cases} 0 & \text{if } x = 0, \\ +\infty & \text{otherwise.} \end{cases}$$

The above large deviation statement therefore gives a concentration of  $\frac{1}{T}\varsigma_T$  or  $\frac{1}{T}\Delta s_{\mathcal{E},T}$  at zero which is faster than exponential.

For the second case, let us make the following assumption.

---

<sup>1</sup>See Section 4 of [HJPR18] for the proof.

**TRI** We say that a repeated interaction system satisfying [ADRI](#) satisfies *time-reversal invariance* if for every  $s \in [0, 1]$  there exist two antiunitary involutions  $C_S(s) : \mathcal{H}_S \rightarrow \mathcal{H}_S$  and  $C_E(s) : \mathcal{H}_E \rightarrow \mathcal{H}_E$  such that if  $C(s) = C_S(s) \otimes C_E(s)$  one has for all  $s \in [0, 1]$

$$[h_S, C_S(s)] = 0, \quad [h_E(s), C_E(s)] = 0, \quad [v(s), C(s)] = 0.$$

This holds for example if each  $h_S, h_E$  and  $v$  are real valued matrices in the same basis, and  $C_S, C_E$  are complex conjugation in the corresponding basis.

In this case, one finds that for all  $s \in [0, 1]$ ,

$$\lambda^{(\alpha)}(s) = \lambda^{(-1-\alpha)}(s), \quad (12.24)$$

or equivalently that the function  $\alpha \mapsto \lambda^{(\alpha)}(s)$  is symmetric about  $\alpha = -1/2$  for all  $s \in [0, 1]$ . This is shown in Lemma 3.11 of [HJPR18](#).

The symmetry (12.24) is central to the Gallavotti–Cohen Theorem which relates in a parameter free formulation the probabilities of observing entropies of opposite signs. Indeed, it implies that  $\Lambda$  is symmetric about the  $\alpha = -1/2$  axis. A direct computation shows that

$$\Lambda^*(x) = x + \Lambda^*(-x). \quad (12.25)$$

A consequence of Theorem 12.8.3 together with this equality is that if e.g.  $\Lambda''(0) \neq 0$ ,

$$\lim_{\delta \rightarrow 0} \lim_{T \rightarrow \infty} \frac{1}{T} \log \frac{\mathbb{P}_T^F(+\Delta s_{E,T} \in [s - \delta, s + \delta])}{\mathbb{P}_T^F(-\Delta s_{E,T} \in [s - \delta, s + \delta])} = -s. \quad (12.26)$$

Roughly speaking, the probability of obtaining a sequence of measurements corresponding to an increase in the entropy of the chain is exponentially smaller than the probability of observing a decrease in entropy. In light of the balance equation (12.12), since  $\Delta s_{S,T}(\omega)$  is bounded as  $T \rightarrow \infty$  (i.e. the finite-size system can only add or remove a finite amount of entropy), this corresponds to an exponentially larger probability for  $\zeta_T$  to be positive than negative. This is a more refined statement than the fact that  $\mathbb{E}[\zeta_T] = \sigma_T^{\text{tot}} \geq 0$ . (12.26) is obtained by observing that in the present case,  $\Lambda$  is analytic in a neighbourhood of the real axis, and so is  $\Lambda^*$  if  $\Lambda$  is strictly convex. See e.g. [EHM09](#); [JOPP12](#); [CJPS18](#) for more information on the role of symmetries such as (12.25).

The next consequence of Theorem 12.8.3 is a result similar to a law of large numbers for  $\Delta y_T^{\text{tot}}$ :

**Corollary 12.8.4.** *Under the same assumptions as in Theorem 12.8.3, for all  $\epsilon > 0$  there exists  $r_\epsilon > 0$  such that for  $T$  large enough*

$$\mathbb{P}_T^F(|\frac{1}{T} \Delta y_T^{\text{tot}} - \Lambda'_Y(0)| > \epsilon) \leq \exp(-r_\epsilon T). \quad (12.27)$$

*Remark.* Such a result is sometimes called exponential convergence. If one could replace  $\mathbb{P}_T^F$  by a  $T$ -independent probability measure  $\mathbb{P}^F$  in (12.27) then the Borel–Cantelli lemma would imply that  $\frac{1}{T} \Delta y_T^{\text{tot}}$  converges  $\mathbb{P}^F$ -almost-surely to  $\Lambda'_Y(0)$ . It implies, however, that  $\lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}(\Delta y_T^{\text{tot}}) = \Lambda'_Y(0)$ . In the case  $Y = \beta h_\mathcal{E}$ , the positivity of  $\sigma_T^{\text{tot}}$  implies  $\Lambda'(0) \geq 0$ . The formula

$$\Lambda'_Y(0) = \int_0^1 \beta(s) \operatorname{tr} \left( X(s) (\operatorname{id} \otimes h_\mathcal{E}(s)) \right) ds,$$

valid for  $Y = \beta h_\mathcal{E}$ , shows that  $\Lambda'(0) = 0$  if  $X(s) \equiv 0$ . This expression can be rewritten in a more illuminating way [BCJP], using the following relation in which the dependence on  $s$  is suppressed:

$$\begin{aligned} D(\rho^{\text{inv}} \otimes \xi \| U^*(\rho^{\text{inv}} \otimes \xi)U) &= -S(\rho^{\text{inv}} \otimes \xi) - \operatorname{tr}[(\rho^{\text{inv}} \otimes \xi) \log(U^* \rho^{\text{inv}} \otimes \xi U)] \\ &= -S(\rho^{\text{inv}}) - S(\xi) - \operatorname{tr}[U(\rho^{\text{inv}} \otimes \xi)U^* \log(\rho^{\text{inv}} \otimes \xi)] \\ &= -S(\rho^{\text{inv}}) - S(\xi) - \operatorname{tr}[U(\rho^{\text{inv}} \otimes \xi)U^* (\log(\rho^{\text{inv}}) \otimes \mathbf{1})] \\ &\quad - \operatorname{tr}[U(\rho^{\text{inv}} \otimes \xi)U^* (\mathbf{1} \otimes \log(\xi))] \\ &= -S(\rho^{\text{inv}}) - S(\xi) + S(\rho^{\text{inv}}) - \operatorname{tr}[U(\rho^{\text{inv}} \otimes \xi)U^* (\mathbf{1} \otimes \log(\xi))] \\ &= \operatorname{tr}[(\xi - \xi^f) \log \xi] = \beta \operatorname{tr}[(\xi^f - \xi) h_\mathcal{E}] \end{aligned}$$

where  $\xi^f := \operatorname{tr}_\mathcal{S}[U(\rho^{\text{inv}} \otimes \xi)U^*]$ . Hence, restoring the dependence on  $s$ ,

$$\beta(s) \operatorname{tr} \left( X(s) (\operatorname{id} \otimes h_\mathcal{E}(s)) \right) = D\left(\rho^{\text{inv}}(s) \otimes \xi(s) \left\| U^*(s)(\rho^{\text{inv}}(s) \otimes \xi(s))U(s) \right.\right)$$

and thus

$$\Lambda'_Y(0) = \int_0^1 D\left(\rho^{\text{inv}}(s) \otimes \xi(s) \left\| U^*(s)(\rho^{\text{inv}}(s) \otimes \xi(s))U(s) \right.\right) ds.$$

By the invariance of the relative entropy under joint unitary conjugations, we may rewrite this formula more naturally as

$$\Lambda'_Y(0) = \int_0^1 D\left(U(s)(\rho^{\text{inv}}(s) \otimes \xi(s))U^*(s) \left\| \rho^{\text{inv}}(s) \otimes \xi(s) \right.\right) ds. \quad (12.28)$$

In particular,  $\Lambda'_Y(0) = 0$  if and only if  $U^*(s)(\rho^{\text{inv}}(s) \otimes \xi(s))U(s) = \rho^{\text{inv}}(s) \otimes \xi(s)$  for all  $s \in [0, 1]$ , and hence  $X(s) = 0$  if and only if  $\Lambda'_Y(0) = 0$ .

One can also obtain a central limit-type result by a slight improvement of the results in Theorem 12.8.3.

**Theorem 12.8.5.** *Under the same assumptions as in Theorem 12.8.3 we have*

$$\frac{1}{\sqrt{T}}(\Delta y_T^{\text{tot}} - T \Lambda_Y'(0)) \xrightarrow{T \rightarrow \infty} \mathcal{N}(0, \Lambda_Y''(0))$$

*in distribution.*

*Remark.* In the case  $Y = \beta h_{\mathcal{E}}$ , Theorem 12.8.5 and the remark following Corollary 12.8.4 show that if  $\Lambda''(0) \neq 0$  (which is generically expected) then  $\sigma_T^{\text{tot}} \rightarrow \infty$  as  $T \rightarrow \infty$ .

## 12.9 Examples

Let us recall the simplest non-trivial RIS, which was already considered in Example 11.2.1. Here, we change notation to call the reduced dynamics  $\mathcal{L}$  instead of  $\Phi$ . In this example, the system and probes are 2-level systems, with  $\mathcal{H}_{\mathcal{S}} = \mathcal{H}_{\mathcal{E}} = \mathbb{C}^2$ , and we choose Hamiltonians  $h_{\mathcal{S}} := E a^* a$  and  $h_{\mathcal{E}_k} \equiv h_{\mathcal{E}} := E_0 b^* b$  where  $a/a^*$  (resp.  $b/b^*$ ) are the Fermionic annihilation/creation operators for  $\mathcal{S}$  (resp.  $\mathcal{E}$ ), with  $E, E_0 > 0$  constants with units of energy. As matrices in the (ground state, excited state) bases  $\{|0\rangle, |1\rangle\}$  for  $\mathcal{S}$  and  $\mathcal{E}$ , we write

$$a = b = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad a^* = b^* = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad a^* a = b^* b = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

As in Example 11.2.1, we consider a constant potential  $v_{\text{RW}} \in \mathcal{B}(\mathcal{H}_{\mathcal{S}} \otimes \mathcal{H}_{\mathcal{E}})$ ,

$$v_{\text{RW}} = \frac{\mu_1}{2}(a^* \otimes b + a \otimes b^*)$$

where  $\mu_1 = 1$  with units of energy. Given  $s \mapsto \beta(s) \in [0, 1]$  a  $C^2$  curve of inverse probe temperatures, an interaction time  $\tau > 0$  and coupling constant  $\lambda > 0$ , we let

$$U = \exp(-i\tau(h_{\mathcal{S}} + h_{\mathcal{E}} + \lambda v_{\text{RW}})).$$

Then  $\text{spec } \mathcal{L}(s)$  is independent of  $s$ , with 1 as a simple eigenvalue with eigenvector

$$\rho^{\text{inv}}(s) = \exp(-\beta^*(s)h_{\mathcal{S}}) / \text{tr}(\exp(-\beta^*(s)h_{\mathcal{S}}))$$

for  $\beta^*(s) = \frac{E_0}{E}\beta(s)$ .

With  $\nu := \sqrt{(E - E_0)^2 + \lambda^2}$ , the assumption  $\nu\tau \notin 2\pi\mathbb{Z}$  yields that  $\mathcal{L}(s)$  is primitive, and moreover, the fact that  $[v_{\text{RW}}, a^*a + b^*b] = 0$  yields  $X(s) \equiv 0$ . Here, we may take  $k_{\mathcal{S}} \equiv \frac{E_0}{E}h_{\mathcal{S}}$  independently of  $s$ , which satisfies  $[k_{\mathcal{S}} + h_{\mathcal{E}}, U] \equiv 0$ .

We choose an initial system state  $\rho^i > 0$ , and set  $Y(s) := \beta(s)h_{\mathcal{E}}$ ,  $A^i := \log \rho^i$ ,  $A^f := \log \rho_T^f$ , for  $\rho_T^f := \mathcal{L}(\frac{T}{\tau}) \cdots \mathcal{L}(\frac{1}{\tau})\rho^i$ . Then Theorem 12.8.1 yields the asymptotic moment generating function of  $\varsigma_T$ : if  $\rho^i$  has spectral decomposition  $\rho^i = r_0|v_0\rangle\langle v_0| + r_1|v_1\rangle\langle v_1| > 0$ , then

$$\begin{aligned} \lim_{T \rightarrow \infty} M_{\varsigma_T}(\alpha) &= (1 + e^{-\beta(0)E_0})^\alpha \\ &\quad \left( r_0^{1+\alpha} (|\langle 0|v_0\rangle|^2 + |\langle 1|v_0\rangle|^2 e^{\alpha\beta(0)E_0}) \alpha \right. \\ &\quad \left. + r_1^{1+\alpha} (|\langle 0|v_1\rangle|^2 + |\langle 1|v_1\rangle|^2 e^{\alpha\beta(0)E_0}) \right). \end{aligned} \quad (12.29)$$

Now, let us consider the same example with a different interaction potential. Instead of  $v_{\text{RW}}$ , we consider the so-called full-dipole interaction potential  $v_{\text{FD}} \in \mathcal{B}(\mathcal{H}_{\mathcal{S}} \otimes \mathcal{H}_{\mathcal{E}})$  given by

$$v_{\text{FD}} := \frac{\mu_1}{2}(a + a^*) \otimes (b + b^*),$$

This example was considered in [HJPR17, Section 7.1], where it was shown that [Prim](#) is satisfied, and that  $\sigma_T \rightarrow \infty$  with a finite and nonzero rate  $\lim_{T \rightarrow \infty} \frac{1}{T}\sigma_T$  for generic choices of parameters  $\{E, E_0, \tau\}$ .

We take  $Y(s) = \beta(s)h_{\mathcal{E}}$  before. Introducing the shorthand  $\eta := \sqrt{(E_0 + E)^2 + \lambda^2}$ , we compute a matrix expression for  $\mathcal{L}_s^{(\alpha)}$  by identifying  $\mathcal{B}(\mathcal{H}_{\mathcal{S}}) \cong \text{Mat}_{2 \times 2}(\mathbb{C}) \cong \mathbb{C}^4$  via  $\begin{pmatrix} \eta_{11} & \eta_{12} \\ \eta_{21} & \eta_{22} \end{pmatrix} \mapsto \begin{pmatrix} \eta_{11} \\ \eta_{12} \\ \eta_{21} \\ \eta_{22} \end{pmatrix}$ . Working in the (ground state, excited state) basis for  $\mathcal{S}$ , we obtain

$$\mathcal{L}^{(\alpha)}(s) = \begin{pmatrix} a & 0 & 0 & d \\ 0 & b & c & 0 \\ 0 & c & e & 0 \\ f & 0 & 0 & g \end{pmatrix},$$

where

$$\begin{aligned} a &= \frac{(2(E_0 + E)^2 + \lambda^2 + \lambda^2 \cos(\eta\tau))}{2(1 + e^{E_0\beta(s)})\eta^2 e^{-E_0\beta(s)}} + \frac{2(E_0 - E)^2 + \lambda^2 + \lambda^2 \cos(\nu\tau)}{2(1 + e^{E_0\beta(s)})\nu^2}, \\ d &= \lambda^2 \left( -\frac{2e^{-E_0\alpha\beta(s)}(\cos(\eta\tau) - 1)}{4(1 + e^{E_0\beta(s)})\eta^2} - \frac{2e^{E_0(\alpha+1)\beta(s)}(\cos(\nu\tau) - 1)}{4(1 + e^{E_0\beta(s)})\nu^2} \right), \end{aligned}$$

$$\begin{aligned}
c &= \frac{\lambda^2 \cosh(E_0 \beta(s) (\frac{1}{2} + \alpha)) \operatorname{sech}\left(\frac{E_0 \beta(s)}{2}\right) \sin\left(\frac{\eta \tau}{2}\right) \sin\left(\frac{\nu \tau}{2}\right)}{\sqrt{E_0^4 + 2(\lambda^2 - E^2) E_0^2 + (E^2 + \lambda^2)^2}}, \\
b &= \frac{(i\eta \cos(\frac{\eta \tau}{2}) + (E_0 + E) \sin(\frac{\eta \tau}{2})) ((E_0 - E) \sin(\frac{\nu \tau}{2}) - i\nu \cos(\frac{\nu \tau}{2}))}{\sqrt{E_0^4 + 2(\lambda^2 - E^2) E_0^2 + (E^2 + \lambda^2)^2}}, \\
e &= \frac{(-e^{i\nu \tau} E_0 + E_0 - E + \nu + e^{i\nu \tau} (E + \nu)) (\eta \cos(\frac{\eta \tau}{2}) + i(E_0 + E) \sin(\frac{\eta \tau}{2}))}{2\eta \nu e^{\frac{1}{2} i\nu \tau}}, \\
f &= \frac{e^{-E_0 \alpha \beta(s)} \lambda^2}{4(1 + e^{E_0 \beta(s)})} \left( \frac{2 - 2\cos(\nu \tau)}{\nu^2} - \frac{2e^{E_0(2\alpha+1)\beta(s)} (\cos(\eta \tau) - 1)}{\eta^2} \right), \\
g &= \frac{(2(E_0 + E)^2 + \lambda^2 + \lambda^2 \cos(\eta \tau))}{2(1 + e^{-E_0 \beta(s)}) \eta^2 e^{E_0 \beta(s)}} + \frac{2(E_0 - E)^2 + \lambda^2 + \lambda^2 \cos(\nu \tau)}{2(1 + e^{-E_0 \beta(s)}) \nu^2},
\end{aligned}$$

which depend on  $s$  through  $\beta(s)$ . The computation was performed with Mathematica, using [Han16a]. We make a particular choice of parameters,  $\lambda = 2$ ,  $\tau = 0.5$ ,  $E_0 = 0.8$ ,  $E = 0.9$ , and two choices of  $[0, 1] \ni s \mapsto \beta(s)$ :

$$\beta_1(s) = \frac{2(3 + 4 \tanh(2s))}{3 + 2 \log(\cosh(2))} \quad (12.30)$$

and

$$\beta_2(s) = a_1 \tanh(2s) - a_2 \tanh\left(\frac{s}{2}\right) - a_3 s^3 + a_4 s^2 - a_5 s + a_6 \quad (12.31)$$

for  $a_1 = 35.483$ ,  $a_2 = 141.929$ ,  $a_3 = 42.945$ ,  $a_4 = 93.5$ ,  $a_5 = 17.808$ ,  $a_6 = 1.061$ . We have  $\beta_1(0) = \beta_2(0) = 1.06$ , and  $\beta_1(1) = \beta_2(1) = 2.43$ , as well as  $\int_0^1 \beta_1(s) ds = \int_0^1 \beta_2(s) ds = 2$ . These are plotted in Figure 12.2.

We compute numerically the function  $\Lambda(\alpha)$  for each choice of  $s \mapsto \beta(s)$ , as shown in Figure 12.3. Figures 12.4 and 12.5 shows the convergence described by Theorem 12.8.5 by simulating 2,000 instances of this repeated interaction system at four values of  $T$ .

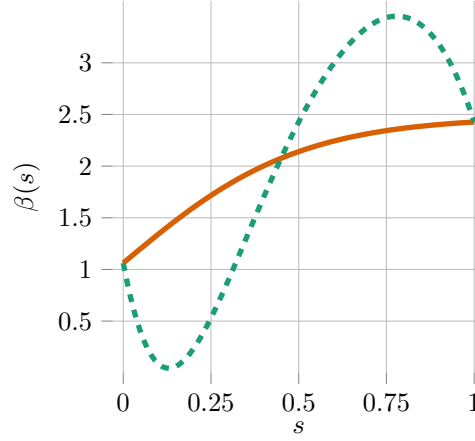


Fig. 12.2 Two choices of curves  $s \mapsto \beta(s)$ . The solid orange line,  $\beta(s) = \beta_1(s)$ , is given by (12.30), and in dashed green line,  $\beta(s) = \beta_2(s)$ , is given by (12.31). These curves of inverse temperatures start and end at the same points, and have the same integral. The differences in  $\Lambda$  and  $\Lambda^*$  shown in Figure 12.3 below are due to the differences in  $\beta_1(s)$  and  $\beta_2(s)$  shown here.

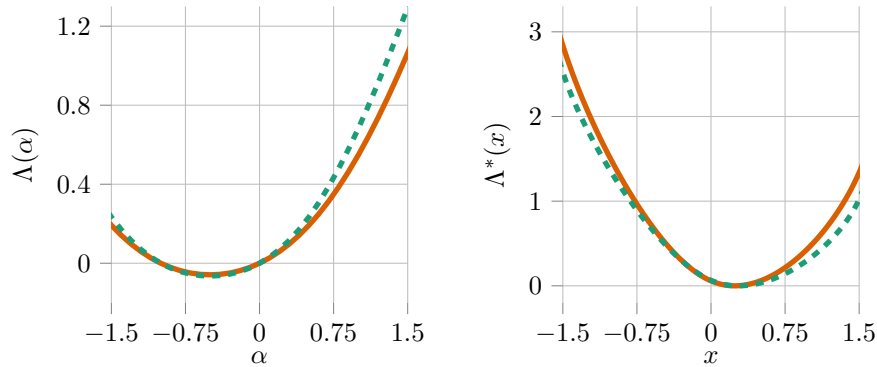


Fig. 12.3 Left: The function  $\Lambda(\alpha)$  for  $Y = \beta h_{\mathcal{E}}$  with  $v_k = v_{\text{FD}}$ ,  $\lambda = 2$ ,  $\tau = 0.5$ ,  $E_0 = 0.8$ ,  $E = 0.9$ , plotted for each choice of  $\beta(s)$ . Right: The rate function  $\Lambda^*(\alpha)$ , for the same setup. In each plot, the solid orange line corresponds to the choice  $\beta(s) = \beta_1(s)$ , defined in (12.30), and the dashed green line corresponds to  $\beta(s) = \beta_2(s)$ , defined in (12.31).



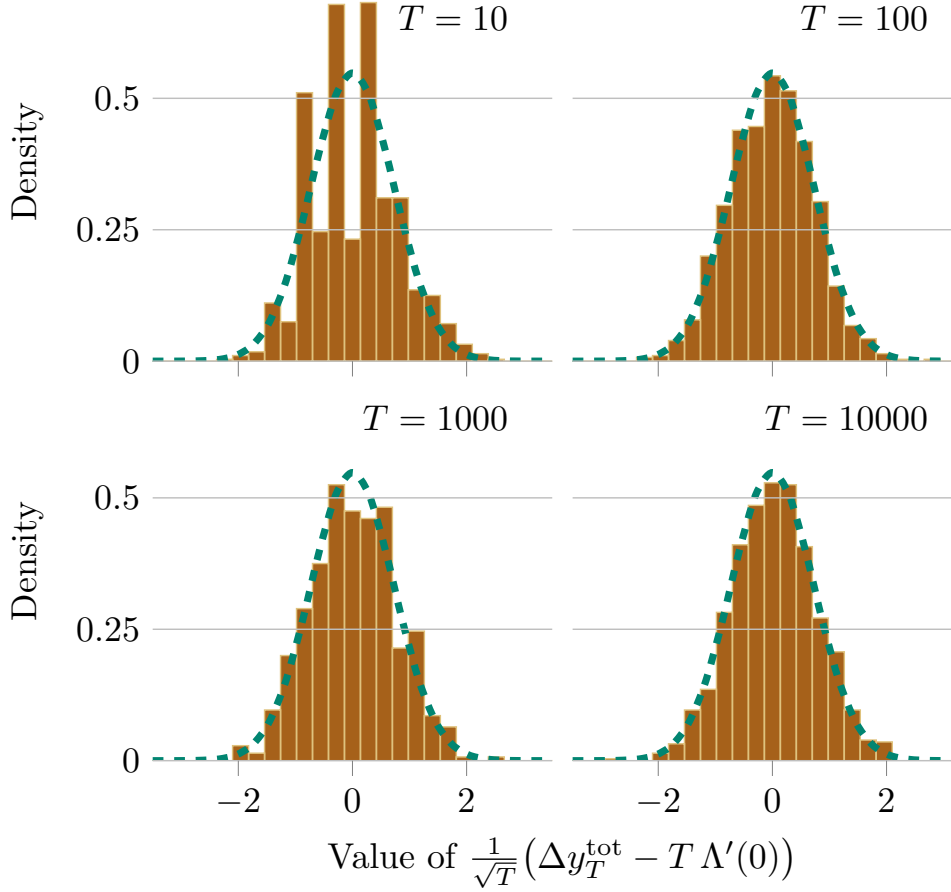


Fig. 12.4 Convergence of  $\frac{1}{\sqrt{T}}(\Delta y_T^{\text{tot}} - T \Lambda'(0))$  to a normal distribution, where  $\Lambda'(0) \approx 0.240$ , with  $\beta(s) = \beta_1(s)$  given by (12.30). Each plot was generated by simulating the two-time measurement protocol in 2,000 instances of the repeated interaction system with the parameters given in Figure 12.3. The value of  $\frac{1}{\sqrt{T}}(\Delta y_T^{\text{tot}} - T \Lambda'(0))$  was calculated for each instance and plotted in a histogram in orange, with bar heights normalized to yield total mass 1. In green, the probability density function of  $\mathcal{N}(0, \Lambda''(0))$  is plotted, where  $\Lambda''(0) \approx 0.530$ . As  $T$  increases, one sees qualitatively the convergence of  $\frac{1}{\sqrt{T}}(\Delta y_T^{\text{tot}} - T \Lambda'(0))$  to the normal distribution, as guaranteed by Theorem 12.8.5.

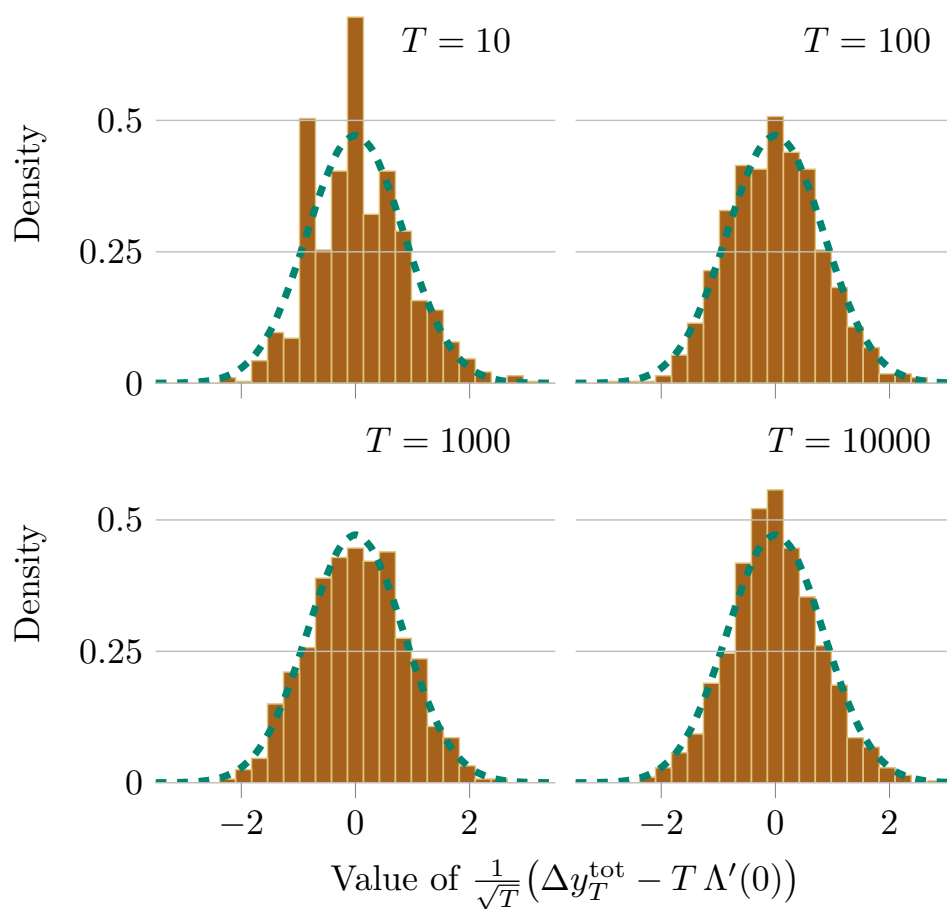


Fig. 12.5 The same setup as Figure 12.4, with  $\beta(s) = \beta_2(s)$  given by (12.31). Here,  $\Lambda'(0) \approx 0.275$ , and  $\Lambda''(0) \approx 0.716$ .

# Chapter 13

## Open questions from Part II

1. In Theorem [11.3.7](#) it is established that faithful PPT quantum channels are eventually entanglement-breaking. Can this result be extended to the non-faithful case? Moreover, can a universal finite number  $n \in \mathbb{N}$  be found such that any PPT quantum channel  $\Phi$  is such that its  $n$ -fold composition  $\Phi^n$  is entanglement-breaking? The PPT<sup>2</sup> conjecture is that such  $n = 2$  suffices (even in the non-trace-preserving case).
2. In Chapter [12](#), a two-time measurement protocol is used to provide a detailed analysis of the entropy production of repeated interaction systems, and has similarly been used in other studies in quantum statistical mechanics. Can such a protocol play a role in an information-theoretic analysis as well?



# References

- [ABL64] Yakir Aharonov, Peter G. Bergmann, and Joel L. Lebowitz. “Time Symmetry in the Quantum Process of Measurement”. *Phys. Rev. (2)* 134 (1964), B1410–B1416.
- [AS19] Mohammad A. Alhejji and Graeme Smith. “A Tight Uniform Continuity Bound for Equivocation” (Sept. 2, 2019). arXiv: [1909.00787 \[quant-ph\]](#).
- [AF04] R. Alicki and M. Fannes. “Continuity of Quantum Conditional Information”. *Journal of Physics A: Mathematical and General* 37.5 (2004), p. L55.
- [Ari96] E. Arikan. “An Inequality on Guessing and Its Application to Sequential Decoding”. *IEEE Transactions on Information Theory* 42.1 (Jan. 1996), pp. 99–105. DOI: [10.1109/18.481781](#).
- [Arn07] Barry C. Arnold. “Majorization: Here, There and Everywhere”. *Statistical Science* 22.3 (Aug. 2007), pp. 407–413. DOI: [10.1214/0883423060000000097](#).
- [AHD19] Koenraad Audenaert, Eric P. Hanson, and Nilanjana Datta. “A Continuity Bound for the Expected Number of Connected Components of a Random Graph: A Model for Epidemics” (Dec. 11, 2019). arXiv: [1912.05599 \[math\]](#).
- [Aud07] Koenraad M. R. Audenaert. “A Sharp Continuity Estimate for the von Neumann Entropy”. *Journal of Physics A: Mathematical and Theoretical* 40.28 (2007), p. 8127. DOI: [10.1088/1751-8113/40/28/S18](#).
- [Bae11] John C. Baez. “Renyi Entropy and Free Energy” (Feb. 10, 2011). arXiv: [1102.2098 \[quant-ph\]](#).
- [BCHW15] Stefan Bäuml, Matthias Christandl, Karol Horodecki, and Andreas Winter. “Limitations on Quantum Key Repeaters”. *Nature Communications* 6 (Apr. 23, 2015), p. 6908. DOI: [10.1038/ncomms7908](#).
- [BS19a] Luis Benet and David P. Sanders. *JuliaIntervals/IntervalArithmetic.jl*. Version 0.16.1. 2014–2019. URL: <https://github.com/JuliaIntervals/IntervalArithmetic.jl/>.
- [BS19b] Luis Benet and David P. Sanders. *JuliaIntervals/IntervalRootFinding.jl*. Version 0.5.1. 2014–2019. URL: <https://github.com/JuliaIntervals/IntervalRootFinding.jl>.
- [BJPP18] T. Benoist, V. Jakšić, Y. Pautrat, and C.-A. Pillet. “On Entropy Production of Repeated Quantum Measurements I. General Theory”. *Communications in Mathematical Physics* 357.1 (Jan. 2018), pp. 77–123. DOI: [10.1007/s00220-017-2947-1](#).

- [BCJP] Tristan Benoist, Noé Cuneo, Vojkan Jakšić, and Claude-Alain Pillet. *On Entropy Production of Repeated Quantum Measurements II. Examples*. In preparation.
- [BFJP17] Tristan Benoist, Martin Fraas, Vojkan Jakšić, and Claude-Alain Pillet. “Full Statistics of Erasure Processes: Isothermal Adiabatic Theory and a Statistical Landauer Principle”. *Rev. Roumaine Math. Pures Appl* 62 (2017), pp. 259–286.
- [BKHW12] *Banff International Research Station Workshop: Operator Structures in Quantum Information Theory*. Ed. by M. Beth Ruskai, D. Kribs, P. Hayden, and A. Winter. 2012. URL: <https://www.birs.ca/workshops/2012/12w5084/report12w5084.pdf>.
- [BEKS17] J. Bezanson, A. Edelman, S. Karpinski, and V. Shah. “Julia: A Fresh Approach to Numerical Computing”. *SIAM Review* 59.1 (Jan. 1, 2017), pp. 65–98. DOI: [10.1137/141000671](https://doi.org/10.1137/141000671).
- [Bha97] Rajendra Bhatia. *Matrix Analysis*. Vol. 169. Graduate Texts in Mathematics. New York, NY: Springer New York, 1997. ISBN: 978-1-4612-6857-4 978-1-4612-0653-8. DOI: [10.1007/978-1-4612-0653-8](https://doi.org/10.1007/978-1-4612-0653-8).
- [BDNK19] Siddharth Bhela, Deepjyoti Deka, Harsha Nagarajan, and Vassilis Kekatos. “Designing Power Grid Topologies for Minimizing Network Disturbances: An Exact MILP Formulation” (Mar. 20, 2019). arXiv: [1903.08354](https://arxiv.org/abs/1903.08354) [math].
- [BD15] Peter J. Bickel and Kjell A. Doksum. *Mathematical Statistics—Basic Ideas and Selected Topics. Vol. 1*. 2nd ed. Texts in Statistical Science Series. CRC Press, Boca Raton, FL, 2015. xxiv+556. ISBN: 978-1-4987-2380-0.
- [Bil95] Patrick Billingsley. *Probability and Measure*. 3rd ed. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1995. xiv+593. ISBN: 0-471-00710-2.
- [BBHF+19] G. M. Bosyk, G. Bellomo, F. Holik, H. Freytes, and G. Sergioli. “Optimal Common Resource in Majorization-Based Resource Theories” (Feb. 5, 2019). arXiv: [1902.01836](https://arxiv.org/abs/1902.01836) [quant-ph].
- [BSFH+17] G. M. Bosyk, G. Sergioli, H. Freytes, F. Holik, and G. Bellomo. “Approximate Transformations of Bipartite Pure-State Entanglement from the Majorization Lattice”. *Physica A: Statistical Mechanics and its Applications* 473 (May 2017), pp. 403–411. DOI: [10.1016/j.physa.2016.12.083](https://doi.org/10.1016/j.physa.2016.12.083). arXiv: [1608.04818](https://arxiv.org/abs/1608.04818).
- [BZHP+16] G. M. Bosyk, S. Zozor, F. Holik, M. Portesi, and P. W. Lamberti. “A Family of Generalized Quantum Entropies: Definition and Properties”. *Quantum Information Processing* 15.8 (Aug. 1, 2016), pp. 3393–3420. DOI: [10.1007/s11128-016-1329-5](https://doi.org/10.1007/s11128-016-1329-5).
- [BV04] Stephen P. Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge, UK ; New York: Cambridge University Press, 2004. 716 pp. ISBN: 978-0-521-83378-3.
- [BDDA+99] V. Bužek, G. Drobný, R. Derka, G. Adam, and H. Wiedemann. “Quantum State Reconstruction from Incomplete Data”. *Chaos, Solitons & Fractals* 10.6 (1999), pp. 981–1074. DOI: [10.1016/S0960-0779\(98\)00144-1](https://doi.org/10.1016/S0960-0779(98)00144-1).

- [CP16] Raffaella Carbone and Yan Pautrat. “Irreducible Decompositions and Stationary States of Quantum Channels”. *Reports on Mathematical Physics* 77.3 (June 2016), pp. 293–313. DOI: [10.1016/S0034-4877\(16\)30032-5](https://doi.org/10.1016/S0034-4877(16)30032-5). arXiv: [1507.08404](https://arxiv.org/abs/1507.08404).
- [Car16] Daniel Cariello. “Completely Reducible Maps in Quantum Information Theory”. *IEEE Transactions on Information Theory* 62.4 (2016), pp. 1721–1732. DOI: [10.1109/TIT.2016.2535176](https://doi.org/10.1109/TIT.2016.2535176).
- [Car19] Daniel Cariello. “Sinkhorn–Knopp Theorem for PPT States”. *Letters in Mathematical Physics* (2019). DOI: [10.1007/s11005-019-01169-9](https://doi.org/10.1007/s11005-019-01169-9).
- [CW03] Kai Chen and Ling-An Wu. “A Matrix Realignment Method for Recognizing Entanglement”. *Quantum Info. Comput.* 3.3 (May 2003), pp. 193–202.
- [CCWF15] Weien Chen, Yongzhi Cao, Hanpin Wang, and Yuan Feng. “Minimum Guesswork Discrimination between Quantum States”. *Quantum Information & Computation* 15.9-10 (July 1, 2015), pp. 737–758.
- [CMNF17] ZhiHua Chen, ZhiHao Ma, Ismail Nikoufar, and Shao-Ming Fei. “Sharp Continuity Bounds for Entropy and Conditional Entropy”. *Science China Physics, Mechanics & Astronomy* 60.2, 020321 (Feb. 2017). DOI: [10.1007/s11433-016-0367-x](https://doi.org/10.1007/s11433-016-0367-x). arXiv: [1701.02398](https://arxiv.org/abs/1701.02398) [[quant-ph](#)].
- [CHDH18a] Hao-Chung Cheng, Eric P. Hanson, Nilanjana Datta, and Min-Hsiu Hsieh. “Duality between Source Coding with Quantum Side Information and C-q Channel Coding” (Sept. 28, 2018). arXiv: [1809.11143](https://arxiv.org/abs/1809.11143) [[quant-ph](#)].
- [CHDH18b] Hao-Chung Cheng, Eric P. Hanson, Nilanjana Datta, and Min-Hsiu Hsieh. “Non-Asymptotic Classical Data Compression with Quantum Side Information” (Mar. 20, 2018). arXiv: [1803.07505](https://arxiv.org/abs/1803.07505) [[quant-ph](#)].
- [CLMO+14] Eric Chitambar, Debbie Leung, Laura Mančinska, Maris Ozols, and Andreas Winter. “Everything You Always Wanted to Know About LOCC (But Were Afraid to Ask)”. *Communications in Mathematical Physics* 328.1 (May 2014), pp. 303–326. DOI: [10.1007/s00220-014-1953-9](https://doi.org/10.1007/s00220-014-1953-9).
- [CF17] Matthias Christandl and Roberto Ferrara. “Private States, Quantum Data Hiding, and the Swapping of Perfect Secrecy”. *Physical Review Letters* 119.22 (Nov. 30, 2017), p. 220506. DOI: [10.1103/PhysRevLett.119.220506](https://doi.org/10.1103/PhysRevLett.119.220506).
- [CV02] F. Cicalese and U. Vaccaro. “Supermodularity and Subadditivity Properties of the Entropy on the Majorization Lattice”. *IEEE Transactions on Information Theory* 48.4 (Apr. 2002), pp. 933–938. DOI: [10.1109/18.992785](https://doi.org/10.1109/18.992785).
- [CGV13] Ferdinando Cicalese, Luisa Gargano, and Ugo Vaccaro. “Information Theoretic Measures of Distances and Their Econometric Applications”. *2013 IEEE International Symposium on Information Theory*. 2013 IEEE International Symposium on Information Theory (ISIT). Istanbul, Turkey: IEEE, July 2013, pp. 409–413. ISBN: 978-1-4799-0446-4. DOI: [10.1109/ISIT.2013.6620258](https://doi.org/10.1109/ISIT.2013.6620258).
- [CLV18] Chris Coey, Miles Lubin, and Juan Pablo Vielma. “Outer Approximation With Conic Certificates For Mixed-Integer Convex Problems” (Aug. 15, 2018). arXiv: [1808.05290](https://arxiv.org/abs/1808.05290) [[math](#)].

- [CT06] Thomas M Cover and Joy A Thomas. *Elements of Information Theory*. 2006. 774 pp.
- [CJPS18] Noé Cuneo, Vojkan Jakšić, Claude-Alain Pillet, and Armen Shirikyan. “Fluctuation Theorem and Thermodynamic Formalism” (Feb. 2, 2018). arXiv: [1712.05167 \[cond-mat, physics:math-ph\]](#).
- [dSLP11] Marcus P. da Silva, Olivier Landon-Cardinal, and David Poulin. “Practical Characterization of Quantum Devices without Tomography”. *Phys. Rev. Lett.* 107.21 (Nov. 2011), p. 210404. DOI: [10.1103/PhysRevLett.107.210404](#).
- [Ell06] Richard S. Ellis. *Entropy, Large Deviations, and Statistical Mechanics*. Classics in Mathematics. Berlin ; New York: Springer, 2006. 364 pp. ISBN: 978-3-540-29059-9.
- [EHM09] Massimiliano Esposito, Upendra Harbola, and Shaul Mukamel. “Nonequilibrium Fluctuations, Fluctuation Theorems, and Counting Statistics in Quantum Systems”. *Reviews of Modern Physics* 81.4 (Dec. 2, 2009), pp. 1665–1702. DOI: [10.1103/RevModPhys.81.1665](#).
- [EH78] David E. Evans and Raphael Høegh-Krohn. “Spectral Properties of Positive Maps on C\*-Algebras”. *Journal of the London Mathematical Society* s2-17.2 (Apr. 1, 1978), pp. 345–355. DOI: [10.1112/jlms/s2-17.2.345](#).
- [FP09] Franco Fagnola and Rely Pellicer. “Irreducible and Periodic Positive Maps”. *Communications on Stochastic Analysis* 3.3 (Dec. 1, 2009). DOI: [10.31390/cosa.3.3.06](#).
- [Fan73] M. Fannes. “A Continuity Property of the Entropy Density for Spin Lattice Systems”. *Communications in Mathematical Physics* 31.4 (Dec. 1, 1973), pp. 291–294. DOI: [10.1007/BF01646490](#).
- [FB14] Serge Fehr and Stefan Berens. “On the Conditional Rényi Entropy”. *IEEE Transactions on Information Theory* 60.11 (Nov. 2014), pp. 6801–6810. DOI: [10.1109/TIT.2014.2357799](#).
- [FL11] Steven T. Flammia and Yi-Kai Liu. “Direct Fidelity Estimation from Few Pauli Measurements”. *Phys. Rev. Lett.* 106.23 (June 2011), p. 230501. DOI: [10.1103/PhysRevLett.106.230501](#).
- [FHLP+07] Laurent Fousse, Guillaume Hanrot, Vincent Lefèvre, Patrick Pélissier, and Paul Zimmermann. “MPFR: A Multiple-Precision Binary Floating-Point Library with Correct Rounding”. *ACM Trans. Math. Softw.* 33.2 (June 2007). DOI: [10.1145/1236463.1236468](#).
- [FL13] Rupert L. Frank and Elliott H. Lieb. “Monotonicity of a Relative Rényi Entropy”. *Journal of Mathematical Physics* 54.12 (Dec. 2013), p. 122201.
- [FYK07] Shigeru Furuichi, Kenjiro Yanagi, and Ken Kuriyama. “A Generalized Fannes’ Inequality”. *Journal of Inequalities in Pure and Applied Mathematics* 8.1 (2007).



- [GABC+20] Gerald Gamrath, Daniel Anderson, Ksenia Bestuzheva, Wei-Kun Chen, Leon Eifler, Maxime Gasse, Patrick Gemander, Ambros Gleixner, Leona Gottwald, Katrin Halbig, Gregor Hendel, Christopher Hojny, Thorsten Koch, Pierre Le Bodic, Stephen J. Maher, Frederic Matter, Matthias Miltenberger, Erik Mühmer, Benjamin Müller, Marc Pfetsch, Franziska Schlösser, Felipe Serrano, Yuji Shinano, Christine Tawfik, Stefan Vigerske, Fabian Wegscheider, Dieter Weninger, and Jakob Witzig. *The SCIP optimization suite 7.0*. Technical report. Optimization Online, Mar. 2020. URL: [http://www.optimization-online.org/DB\\_HTML/2020/03/7705.html](http://www.optimization-online.org/DB_HTML/2020/03/7705.html).
- [Ger77] I. B. Gertsbakh. “Epidemic Process on a Random Graph: Some Preliminary Results”. *Journal of Applied Probability* 14.03 (Sept. 1977), pp. 427–438. DOI: [10.1017/S0021900200025675](https://doi.org/10.1017/S0021900200025675).
- [GKGB+07] Sébastien Gleyzes, Stefan Kuhr, Christine Guerlin, Julien Bernu, Samuel Deléglise, Ulrich Busk Hoff, Michel Brune, Jean-Michel Raimond, and Serge Haroche. “Quantum Jumps of Light Recording the Birth and Death of a Photon in a Cavity”. *Nature* 446.7133 (Mar. 2007), pp. 297–300. DOI: [10.1038/nature05589](https://doi.org/10.1038/nature05589).
- [Gor41] Robert D. Gordon. “Values of Mills’ Ratio of Area to Bounding Ordinate and of the Normal Probability Integral for Large Values of the Argument”. *The Annals of Mathematical Statistics* 12.3 (Sept. 1941), pp. 364–366. DOI: [10.1214/aoms/1177731721](https://doi.org/10.1214/aoms/1177731721).
- [GLS93] Martin Grötschel, László Lovász, and Alexander Schrijver. *Geometric Algorithms and Combinatorial Optimization*. Vol. 2. Algorithms and Combinatorics. Berlin, Heidelberg: Springer Berlin Heidelberg, 1993. ISBN: 978-3-642-78242-8 978-3-642-78240-4. DOI: [10.1007/978-3-642-78240-4](https://doi.org/10.1007/978-3-642-78240-4).
- [GCGP+17] Giacomo Guarnieri, Steve Campbell, John Goold, Simon Pigeon, Bassano Vacchini, and Mauro Paternostro. “Full Counting Statistics Approach to the Quantum Non-Equilibrium Landauer Bound”. *New Journal of Physics* 19.10 (2017), p. 103038.
- [GB02] Leonid Gurvits and Howard Barnum. “Largest Separable Balls around the Maximally Mixed Bipartite Quantum State”. *Physical Review A* 66.6 (Dec. 12, 2002). DOI: [10.1103/PhysRevA.66.062311](https://doi.org/10.1103/PhysRevA.66.062311). arXiv: [quant-ph/0204159](https://arxiv.org/abs/quant-ph/0204159).
- [Han16a] Eric Hanson. *Full dipole code*. 2016. URL: [https://ericphanson.com/code/RIS\\_full\\_dipole](https://ericphanson.com/code/RIS_full_dipole).
- [Han16b] Eric Hanson. “Landauer’s Principle in Repeated Interaction Systems”. Master’s thesis. McGill University, 2016. URL: [http://digitool.library.mcgill.ca/R?func=dbin-jump-full&object\\_id=141477](http://digitool.library.mcgill.ca/R?func=dbin-jump-full&object_id=141477).
- [Han20a] Eric P. Hanson. *ericphanson/AHD19\_supplementary*. Version v1.0.0. Jan. 2020. DOI: [10.5281/zenodo.3632955](https://doi.org/10.5281/zenodo.3632955). URL: <https://doi.org/10.5281/zenodo.3632955>.
- [Han20b] Eric P. Hanson. *ericphanson/GuessworkQuantumSideInfo.jl: v0.1.2*. Version v0.1.2. Sept. 2020. DOI: [10.5281/zenodo.4018151](https://doi.org/10.5281/zenodo.4018151). URL: <https://doi.org/10.5281/zenodo.4018151>.

- [HD18] Eric P. Hanson and Nilanjana Datta. “Maximum and Minimum Entropy States Yielding Local Continuity Bounds”. *Journal of Mathematical Physics* 59.4 (Apr. 1, 2018), p. 042204. DOI: [10.1063/1.5000120](https://doi.org/10.1063/1.5000120).
- [HD17] Eric P. Hanson and Nilanjana Datta. “Tight Uniform Continuity Bound for a Family of Entropies” (July 13, 2017). arXiv: [1707.04249](https://arxiv.org/abs/1707.04249) [[math-ph](#), [physics:quant-ph](#)].
- [HD19] Eric P. Hanson and Nilanjana Datta. “Universal Proofs of Entropic Continuity Bounds via Majorization Flow” (Dec. 11, 2019). arXiv: [1909.06981](https://arxiv.org/abs/1909.06981) [[quant-ph](#)].
- [HJPR18] Eric P. Hanson, Alain Joye, Yan Pautrat, and Renaud Raquépas. “Landauer’s Principle for Trajectories of Repeated Interaction Systems”. *Annales Henri Poincaré* 19.7 (July 1, 2018), pp. 1939–1991. DOI: [10.1007/s00023-018-0679-1](https://doi.org/10.1007/s00023-018-0679-1).
- [HJPR17] Eric P. Hanson, Alain Joye, Yan Pautrat, and Renaud Raquépas. “Landauer’s Principle in Repeated Interaction Systems”. *Communications in Mathematical Physics* 349.1 (Jan. 1, 2017), pp. 285–327. DOI: [10.1007/s00220-016-2751-3](https://doi.org/10.1007/s00220-016-2751-3).
- [HKDW20] Eric P. Hanson, Vishal Katariya, Nilanjana Datta, and Mark M. Wilde. “Guesswork with Quantum Side Information” (Feb. 3, 2020). arXiv: [2001.03598](https://arxiv.org/abs/2001.03598) [[quant-ph](#)].
- [HRS20] Eric P. Hanson, Cambyse Rouzé, and Daniel Stilck França. “Eventually Entanglement Breaking Markovian Dynamics: Structure and Characteristic Times”. *Annales Henri Poincaré* (Mar. 11, 2020). DOI: [10.1007/s00023-020-00906-4](https://doi.org/10.1007/s00023-020-00906-4).
- [HLP29] H. G. Hardy, J. E. Littlewood, and G. Pólya. “Some Simple Inequalities Satisfied by Convex Functions”. *Messenger Mathematics* 58 (1929), pp. 145–152.
- [HM17] Fumio Hiai and Milan Mosonyi. “Different Quantum F-Divergences and the Reversibility of Quantum Operations”. *Reviews in Mathematical Physics* 29.07 (Aug. 2017), p. 1750023. DOI: [10.1142/S0129055X17500234](https://doi.org/10.1142/S0129055X17500234). arXiv: [1604.03089](https://arxiv.org/abs/1604.03089).
- [HMO07] Fumio Hiai, Milán Mosonyi, and Tomohiro Ogawa. “Large Deviations and Chernoff Bound for Certain Correlated States on a Spin Chain”. *Journal of Mathematical Physics* 48.12 (Dec. 2007), p. 123301. DOI: [10.1063/1.2812417](https://doi.org/10.1063/1.2812417).
- [HY10] S. Ho and R. W. Yeung. “The Interplay Between Entropy and Variational Distance”. *IEEE Transactions on Information Theory* 56.12 (Dec. 2010), pp. 5906–5929. DOI: [10.1109/TIT.2010.2080452](https://doi.org/10.1109/TIT.2010.2080452).
- [Hor54] Alfred Horn. “Doubly Stochastic Matrices and the Diagonal of a Rotation Matrix”. *American Journal of Mathematics* 76.3 (July 1954), p. 620. DOI: [10.2307/2372705](https://doi.org/10.2307/2372705). JSTOR: [2372705?origin=crossref](https://www.jstor.org/stable/2372705).

- [HHHO05] Karol Horodecki, Michał Horodecki, Paweł Horodecki, and Jonathan Oppenheim. “Secure Key from Bound Entanglement”. *Physical Review Letters* 94.16 (Apr. 26, 2005), p. 160502. DOI: [10.1103/PhysRevLett.94.160502](https://doi.org/10.1103/PhysRevLett.94.160502).
- [HOS17] M. Horodecki, J. Oppenheim, and C. Sparaciari. “Approximate Majorization”. *ArXiv e-prints* (June 2017). arXiv: [1706.05264](https://arxiv.org/abs/1706.05264) [[quant-ph](#)].
- [HSR03] Michael Horodecki, Peter W. Shor, and Mary Beth Ruskai. “Entanglement Breaking Channels”. *Reviews in Mathematical Physics* 15.06 (Aug. 1, 2003), pp. 629–641. DOI: [10.1142/S0129055X03001709](https://doi.org/10.1142/S0129055X03001709).
- [HHH96] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. “Separability of Mixed States: Necessary and Sufficient Conditions”. *Physics Letters A* 223.1 (Nov. 25, 1996), pp. 1–8. DOI: [10.1016/S0375-9601\(96\)00706-2](https://doi.org/10.1016/S0375-9601(96)00706-2).
- [HOS18] Michał Horodecki, Jonathan Oppenheim, and Carlo Sparaciari. “Extremal Distributions under Approximate Majorization”. *Journal of Physics A: Mathematical and Theoretical* 51.30 (June 2018), p. 305301. DOI: [10.1088/1751-8121/aac87c](https://doi.org/10.1088/1751-8121/aac87c).
- [HP13] Jordan M. Horowitz and Juan M. R. Parrondo. “Entropy Production along Nonequilibrium Quantum Jump Trajectories”. *New Journal of Physics* 15.8 (2013), p. 085028.
- [HY06] Xinhua Hu and Zhongxing Ye. “Generalized Quantum Entropy”. *Journal of Mathematical Physics* 47.2 (Feb. 1, 2006), p. 023502. DOI: [10.1063/1.2165794](https://doi.org/10.1063/1.2165794).
- [JPW14] Vojkan Jakšić, Claude-Alain Pillet, and Matthias Westrich. “Entropic Fluctuations of Quantum Dynamical Semigroups”. *Journal of Statistical Physics* 154.1-2 (Jan. 2014), pp. 153–187. DOI: [10.1007/s10955-013-0826-5](https://doi.org/10.1007/s10955-013-0826-5). arXiv: [1305.4409](https://arxiv.org/abs/1305.4409).
- [JOPP12] Vojkan Jakšić, Yoshiko Ogata, Yan Pautrat, and Claude-Alain Pillet. “Entropic Fluctuations in Quantum Statistical Mechanics. An Introduction”. *Quantum Theory from Small to Large Scales* (2012), pp. 213–410. arXiv: [1106.3786](https://arxiv.org/abs/1106.3786).
- [JOPS12] Vojkan Jakšić, Yoshiko Ogata, Claude-Alain Pillet, and Robert Seiringer. “Quantum Hypothesis Testing and Non-Equilibrium Statistical Mechanics”. *Reviews in Mathematical Physics* 24.6 (2012), pp. 1230002, 67. DOI: [10.1142/S0129055X12300026](https://doi.org/10.1142/S0129055X12300026).
- [JP14] Vojkan Jakšić and Claude-Alain Pillet. “A Note on the Landauer Principle in Quantum Statistical Mechanics”. *J. Math. Phys.* 55.7 (July 2014), 075210–75210:21.
- [Jay57] E. T. Jaynes. “Information Theory and Statistical Mechanics. II”. *Phys. Rev.* 108.2 (Oct. 1957), pp. 171–190. DOI: [10.1103/PhysRev.108.171](https://doi.org/10.1103/PhysRev.108.171).
- [jVRS+19] johnjforrest, Stefan Vigerske, Ted Ralphs, Haroldo Gambini Santos, Lou Hafer, Bjarni Kristjansson, jpfasano, Edwin Straver, Miles Lubin, rlougee, jpgoncall, h-i-gassmann, and Matthew Saltzman. *Coin-or/Cbc: Version 2.10.3*. Zenodo. June 2019. DOI: [10.5281/zenodo.3246628](https://doi.org/10.5281/zenodo.3246628).

- [KMP17] Matthew Kennedy, Nicholas A. Manor, and Vern I. Paulsen. “Composition of PPT Maps” (Oct. 23, 2017). arXiv: [1710.08475](https://arxiv.org/abs/1710.08475) [[math-ph](#)].
- [KRS09] Robert Koenig, Renato Renner, and Christian Schaffner. “The Operational Meaning of Min- and Max-Entropy”. *IEEE Transactions on Information Theory* 55.9 (Sept. 2009), pp. 4337–4347. DOI: [10.1109/TIT.2009.2025545](https://doi.org/10.1109/TIT.2009.2025545). arXiv: [0807.1338](https://arxiv.org/abs/0807.1338).
- [Kog13] H. Koga. “Characterization of the Smooth Rényi Entropy Using Majorization”. *2013 IEEE Information Theory Workshop (ITW)*. 2013 IEEE Information Theory Workshop (ITW). Sept. 2013, pp. 1–5. DOI: [10.1109/ITW.2013.6691332](https://doi.org/10.1109/ITW.2013.6691332).
- [Kur00] Jorge Kurchan. “A Quantum Fluctuation Theorem”. *arXiv preprint* (2000), cond-mat/0007360.
- [LG16] L. Lami and V. Giovannetti. “Entanglement-Saving Channels”. *Journal of Mathematical Physics* 57.3 (Mar. 2016), p. 032201. DOI: [10.1063/1.4942495](https://doi.org/10.1063/1.4942495).
- [Lan61] Rolf Landauer. “Irreversibility and Heat Generation in the Computing Process”. *IBM Journal of Research and Development* 5 (1961), pp. 183–191.
- [LHL07] Reine Lundin, Thijs Holleboom, and Stefan Lindskog. “On the Relationship between Confidentiality Measures: Entropy and Guesswork.” *Proceedings of the 5th International Workshop on Security in Information Systems*. Funchal, Madeira, Portugal: SciTePress, 2007, pp. 135–144. DOI: [10.5220/0002426901350144](https://doi.org/10.5220/0002426901350144).
- [Mac03] David J. C. MacKay. *Information Theory, Inference, and Learning Algorithms*. Cambridge, UK ; New York: Cambridge University Press, 2003. 628 pp. ISBN: 978-0-521-64298-9.
- [MOA11] Albert W. Marshall, Ingram Olkin, and Barry C. Arnold. *Inequalities: Theory of Majorization and Its Applications*. 2nd ed. Springer Series in Statistics. New York: Springer-Verlag, 2011. ISBN: 978-0-387-40087-7.
- [Mat02] Jiří Matoušek, ed. *Lectures on Discrete Geometry*. Vol. 212. Graduate Texts in Mathematics. New York, NY: Springer New York, 2002. DOI: [10.1007/978-1-4613-0039-7](https://doi.org/10.1007/978-1-4613-0039-7).
- [MDSF+13] Martin Müller-Lennert, Frédéric Dupuis, Oleg Szehr, Serge Fehr, and Marco Tomamichel. “On Quantum Renyi Entropies: A New Generalization and Some Properties”. *Journal of Mathematical Physics* 54.12 (Dec. 2013), p. 122203. DOI: [10.1063/1.4838856](https://doi.org/10.1063/1.4838856). arXiv: [1306.3142](https://arxiv.org/abs/1306.3142).
- [Nie99] M. A. Nielsen. “Conditions for a Class of Entanglement Transformations”. *Physical Review Letters* 83.2 (July 12, 1999), pp. 436–439. DOI: [10.1103/PhysRevLett.83.436](https://doi.org/10.1103/PhysRevLett.83.436).
- [NC09] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2009. ISBN: 978-1-107-00217-3. DOI: [10.1017/cbo9780511976667](https://doi.org/10.1017/cbo9780511976667).
- [OCPB16] B. O’Donoghue, E. Chu, N. Parikh, and S. Boyd. “Conic Optimization via Operator Splitting and Homogeneous Self-Dual Embedding”. *Journal of Optimization Theory and Applications* 169.3 (June 2016), pp. 1042–1068.

- [OH17] Jonathan Oppenheim and Michał Horodecki. *Smoothing Majorization*. Private communication. E-mail. 2017.
- [Par99] K. R. Parthasarathy. “Extremal Decision Rules in Quantum Hypothesis Testing”. *Infinite Dimensional Analysis, Quantum Probability and Related Topics* 02.04 (Dec. 1999), pp. 557–568. DOI: [10.1142/S0219025799000321](https://doi.org/10.1142/S0219025799000321).
- [Pet03] Dénes Petz. “Monotonicity of Quantum Relative Entropy Revisited”. *Reviews in Mathematical Physics* 15.01 (Mar. 1, 2003), pp. 79–91. DOI: [10.1142/S0129055X03001576](https://doi.org/10.1142/S0129055X03001576).
- [Pet08] Dénes Petz. *Quantum Information Theory and Quantum Statistics*. 1st ed. Theoretical and Mathematical Physics. Springer-Verlag Berlin Heidelberg, 2008.
- [Pet85] Dénes Petz. “Quasi-Entropies for States of a von Neumann Algebra”. *Publications of the Research Institute for Mathematical Sciences* 21.4 (Aug. 31, 1985), pp. 787–800. DOI: [10.2977/prims/1195178929](https://doi.org/10.2977/prims/1195178929).
- [PR98] Dénes Petz and Mary Beth Ruskai. “Contraction of Generalized Relative Entropy Under Stochastic Mappings on Matrices”. *Infinite Dimensional Analysis, Quantum Probability and Related Topics* 01.01 (Jan. 1, 1998), pp. 83–89. DOI: [10.1142/S0219025798000077](https://doi.org/10.1142/S0219025798000077).
- [Pey15] Juan Peypouquet. *Convex Optimization in Normed Spaces*. Springer International Publishing, 2015. DOI: [10.1007/978-3-319-13710-0](https://doi.org/10.1007/978-3-319-13710-0).
- [Rag95] G. A. Raggio. “Properties of Q-entropies”. *Journal of Mathematical Physics* 36.9 (Sept. 1, 1995), pp. 4785–4791. DOI: [10.1063/1.530920](https://doi.org/10.1063/1.530920).
- [RJP18] Mizanur Rahaman, Samuel Jaques, and Vern I. Paulsen. “Eventually Entanglement Breaking Maps”. *Journal of Mathematical Physics* 59.6 (June 1, 2018), p. 062201. DOI: [10.1063/1.5024385](https://doi.org/10.1063/1.5024385).
- [Ras02] A. E. Rastegin. “Relative Error of State-Dependent Cloning”. *Physical Review A* 66.4 (Oct. 10, 2002), p. 042304. DOI: [10.1103/PhysRevA.66.042304](https://doi.org/10.1103/PhysRevA.66.042304).
- [Ras06] A. E. Rastegin. “Sine Distance for Quantum States” (Feb. 14, 2006). arXiv: [quant-ph/0602112](https://arxiv.org/abs/quant-ph/0602112).
- [Ras11] Alexey E. Rastegin. “Some General Properties of Unified Entropies”. *Journal of Statistical Physics* 143.6 (June 2, 2011), p. 1120. DOI: [10.1007/s10955-011-0231-x](https://doi.org/10.1007/s10955-011-0231-x). arXiv: [1012.5356](https://arxiv.org/abs/1012.5356) [quant-ph].
- [RT91] Pushpa N. Rathie and Inder Jeet Taneja. “Unified (r, s)-Entropy and Its Bivariate Measures”. *Information Sciences* 54.1 (Mar. 1, 1991), pp. 23–39. DOI: [10.1016/0020-0255\(91\)90043-T](https://doi.org/10.1016/0020-0255(91)90043-T).
- [RW14] David Reeb and Michael M. Wolf. “An Improved Landauer Principle with Finite-Size Corrections”. *New Journal of Physics* 16.10 (Oct. 8, 2014), p. 103011. DOI: [10.1088/1367-2630/16/10/103011](https://doi.org/10.1088/1367-2630/16/10/103011). arXiv: [1306.4352](https://arxiv.org/abs/1306.4352).
- [RS81] M. Reed and B. Simon. *I: Functional Analysis*. Methods of Modern Mathematical Physics. Elsevier Science, 1981. ISBN: 978-0-08-057048-8.



- [RR12] Joseph M. Renes and Renato Renner. “One-Shot Classical Data Compression with Quantum Side Information and the Distillation of Common Randomness or Secret Keys”. *IEEE Transactions on Information Theory* 58.3 (Mar. 2012), pp. 1985–1991. DOI: [10.1109/TIT.2011.2177589](https://doi.org/10.1109/TIT.2011.2177589). arXiv: [1008.0452](https://arxiv.org/abs/1008.0452).
- [Ren05] R. Renner. “Security of Quantum Key Distribution”. ETH Zürich, Dec. 2005. arXiv: [quant-ph/0512258](https://arxiv.org/abs/quant-ph/0512258).
- [Rén61] Alfréd Rényi. “On Measures of Entropy and Information”. Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics. The Regents of the University of California, 1961. URL: <https://projecteuclid.org/euclid.bsmmsp/1200512181>.
- [Roc96] Ralph Tyrell Rockafellar. *Convex Analysis*. Princeton Landmarks in Mathematics and Physics. Princeton University Press, 1996. ISBN: 978-0-691-01586-6 0-691-01586-4.
- [Ros81] Sheldon M. Ross. “A Random Graph”. *Journal of Applied Probability* 18.1 (Mar. 1981), pp. 309–315. DOI: [10.2307/3213194](https://doi.org/10.2307/3213194).
- [RS62] J. Barkley Rosser and Lowell Schoenfeld. “Approximate Formulas for Some Functions of Prime Numbers”. *Illinois Journal of Mathematics* 6.1 (Mar. 1962), pp. 64–94.
- [RBCH+01] Pranaw Rungta, V. Bužek, Carlton M. Caves, M. Hillery, and G. J. Milburn. “Universal State Inversion and Concurrence in Arbitrary Dimensions”. *Physical Review A* 64.4 (Sept. 18, 2001), p. 042315. DOI: [10.1103/PhysRevA.64.042315](https://doi.org/10.1103/PhysRevA.64.042315).
- [RD66] B. Russo and H. A. Dye. “A Note on Unitary Operators in C\*-Algebras”. *Duke Mathematical Journal* 33.2 (June 1966), pp. 413–416. DOI: [10.1215/S0012-7094-66-03346-1](https://doi.org/10.1215/S0012-7094-66-03346-1).
- [SMMP93] M. Salicru, M. L. Menendez, D. Morales, and L. Pardo. “Asymptotic Distribution of  $(h, \varphi)$ -Entropies”. *Communications in Statistics - Theory and Methods* 22.7 (Jan. 1, 1993), pp. 2015–2031. DOI: [10.1080/03610929308831131](https://doi.org/10.1080/03610929308831131).
- [Sas13] Igal Sason. “Entropy Bounds for Discrete Random Variables via Maximal Coupling”. *IEEE Transactions on Information Theory* 59.11 (Nov. 2013), pp. 7118–7131. DOI: [10.1109/TIT.2013.2274515](https://doi.org/10.1109/TIT.2013.2274515). arXiv: [1209.5259](https://arxiv.org/abs/1209.5259).
- [Sch96] Benjamin Schumacher. “Sending Entanglement through Noisy Quantum Channels”. *Phys. Rev. A* 54.4 (Oct. 1996), pp. 2614–2628. DOI: [10.1103/PhysRevA.54.2614](https://doi.org/10.1103/PhysRevA.54.2614).
- [Sch23] Issai Schur. “Über Eine Klasse von Mittelbildungen Mit Anwendungen Auf Die Determinantentheorie”. *Sitzungsberichte der Berliner Mathematischen Gesellschaft* 22 (1923), pp. 9–20.
- [Sim15] Barry Simon. *Real Analysis*. A Comprehensive Course in Analysis part 1. Providence, Rhode Island: American Mathematical Society, 2015. 789 pp. ISBN: 978-1-4704-1099-5.

- [Sko16] Maciej Skorski. “How to Smooth Entropy?” *Proceedings of the 42Nd International Conference on SOFSEM 2016: Theory and Practice of Computer Science - Volume 9587*. New York, NY, USA: Springer-Verlag New York, Inc., 2016, pp. 392–403. DOI: [10.1007/978-3-662-49192-8\\_32](https://doi.org/10.1007/978-3-662-49192-8_32).
- [SS12] Łukasz Skowronek and Erling Størmer. “Choi Matrices, Norms and Entanglement Associated with Positive Maps on Matrix Algebras”. *Journal of Functional Analysis* 262.2 (Jan. 2012), pp. 639–647. DOI: [10.1016/j.jfa.2011.09.022](https://doi.org/10.1016/j.jfa.2011.09.022).
- [Sti87] Eberhard Stickel. “On the Fréchet Derivative of Matrix Functions”. *Linear Algebra and its Applications* 91 (1987), pp. 83–88. DOI: [http://dx.doi.org/10.1016/0024-3795\(87\)90062-0](https://dx.doi.org/10.1016/0024-3795(87)90062-0).
- [Tas00] Hal Tasaki. “Jarzynski Relations for Quantum Systems and Some Applications”. *arXiv preprint* (2000), cond-mat/0009244.
- [Teh20] Alireza Tehrani. “The Critical Points of Coherent Information on the Manifold of Positive Definite Matrices”. Master’s thesis. Guelph, Ontario, Canada: The University of Guelph, Jan. 2020. 130 pp.
- [Tom12] Marco Tomamichel. “A Framework for Non-Asymptotic Quantum Information Theory”. Mar. 9, 2012. arXiv: [1203.2142](https://arxiv.org/abs/1203.2142).
- [Tom16] Marco Tomamichel. *Quantum Information Processing with Finite Resources*. Vol. 5. 2016. arXiv: [1504.00233](https://arxiv.org/abs/1504.00233). URL: <http://arxiv.org/abs/1504.00233>.
- [Tsa88] Constantino Tsallis. “Possible Generalization of Boltzmann-Gibbs Statistics”. *Journal of Statistical Physics* 52.1 (July 1, 1988), pp. 479–487. DOI: [10.1007/BF01016429](https://doi.org/10.1007/BF01016429).
- [Tuc11] Warwick Tucker. *Validated Numerics: A Short Introduction to Rigorous Computations*. Princeton: Princeton University Press, 2011. 138 pp. ISBN: 978-0-691-14781-9.
- [UMZH+14] Madeleine Udell, Karanveer Mohan, David Zeng, Jenny Hong, Steven Diamond, and Stephen Boyd. “Convex Optimization in Julia”. *Proceedings of the 1st First Workshop for High Performance Technical Computing in Dynamic Languages*. IEEE Press. 2014, pp. 18–28.
- [vdMee16] Remco van der Meer. “The Properties of Thermodynamical Operations”. Bachelor’s thesis. Delft University of Technology, July 2016.
- [vdMNW17] Remco van der Meer, Nelly Huei Ying Ng, and Stephanie Wehner. “Smoothed Generalized Free Energies for Thermodynamics”. *Physical Review A* 96.6 (Dec. 29, 2017), p. 062135. DOI: [10.1103/PhysRevA.96.062135](https://doi.org/10.1103/PhysRevA.96.062135). arXiv: [1706.03193](https://arxiv.org/abs/1706.03193) [quant-ph].
- [vDH02] W. van Dam and P. Hayden. “Renyi-Entropic Bounds on Quantum Communication”. *eprint arXiv:quant-ph/0204093* (Apr. 2002). eprint: [quant-ph/0204093](https://arxiv.org/abs/quant-ph/0204093).
- [VJN00] Guifré Vidal, Daniel Jonathan, and M. A. Nielsen. “Approximate Transformations and Robust Manipulation of Bipartite Pure-State Entanglement”. *Physical Review A* 62.1 (June 12, 2000), p. 012304. DOI: [10.1103/PhysRevA.62.012304](https://doi.org/10.1103/PhysRevA.62.012304).

- [Vil09] Cédric Villani. *Optimal Transport: Old and New*. Grundlehren Der Mathematischen Wissenschaften. Berlin Heidelberg: Springer-Verlag, 2009. ISBN: 978-3-540-71049-3.
- [WMVF16] Yu-Xin Wang, Liang-Zhu Mu, Vlatko Vedral, and Heng Fan. “Entanglement Rényi  $\alpha$  Entropy”. *Phys. Rev. A* 93.2 (Feb. 2016), p. 022324. DOI: [10.1103/PhysRevA.93.022324](https://doi.org/10.1103/PhysRevA.93.022324).
- [Wat09] John Watrous. “Semidefinite Programs for Completely Bounded Norms”. *Theory of Computing* 5 (Nov. 2009), pp. 217–238. DOI: [10.4086/toc.2009.v005a011](https://doi.org/10.4086/toc.2009.v005a011). arXiv: [0901.4709](https://arxiv.org/abs/0901.4709).
- [Wat18] John Watrous. *The Theory of Quantum Information*. 1st ed. Cambridge University Press, Apr. 26, 2018. ISBN: 978-1-316-84814-2 978-1-107-18056-7. DOI: [10.1017/9781316848142](https://doi.org/10.1017/9781316848142).
- [Wil20] Mark M. Wilde. “Optimal uniform continuity bound for conditional entropy of classical–quantum states”. *Quantum Information Processing* 19.2 (Jan. 2, 2020), p. 61. ISSN: 1573-1332. DOI: [10.1007/s11128-019-2563-4](https://doi.org/10.1007/s11128-019-2563-4).
- [WWY14] Mark M. Wilde, Andreas Winter, and Dong Yang. “Strong Converse for the Classical Capacity of Entanglement-Breaking and Hadamard Channels via a Sandwiched Renyi Relative Entropy”. *Communications in Mathematical Physics* 331.2 (Oct. 2014), pp. 593–622. DOI: [10.1007/s00220-014-2122-x](https://doi.org/10.1007/s00220-014-2122-x). arXiv: [1306.1586](https://arxiv.org/abs/1306.1586).
- [WS17] Matthew Wilhelm and Matthew D. Stuber. “Easy Advanced Global Optimization (EAGO): An Open-Source Platform for Robust and Global Optimization in Julia”. *2017 AIChE Annual Meeting*. AIChE, Oct. 2017. URL: <https://aiche.confex.com/aiche/2017/meetingapp.cgi/Paper/490467>.
- [Win16] Andreas Winter. “Tight Uniform Continuity Bounds for Quantum Entropies: Conditional Entropy, Relative Entropy Distance and Energy Constraints”. *Communications in Mathematical Physics* 347.1 (Oct. 2016), pp. 291–313. DOI: [10.1007/s00220-016-2609-8](https://doi.org/10.1007/s00220-016-2609-8). arXiv: [1507.07775](https://arxiv.org/abs/1507.07775).
- [Wol12] Michael Wolf. “Quantum Channels & Operations Guided Tour”. 2012. URL: <https://www-m5.ma.tum.de/foswiki/pub/M5/Allgemeines/MichaelWolf/QChannelLecture.pdf>.
- [WC08] Michael M. Wolf and J. Ignacio Cirac. “Dividing Quantum Channels”. *Communications in Mathematical Physics* 279.1 (Apr. 1, 2008), pp. 147–168. DOI: [10.1007/s00220-008-0411-y](https://doi.org/10.1007/s00220-008-0411-y).
- [WP10] Michael M. Wolf and David Perez-Garcia. “The Inverse Eigenvalue Problem for Quantum Channels” (May 25, 2010). arXiv: [1005.4545](https://arxiv.org/abs/1005.4545) [[math-ph](#), [physics:quant-ph](#)].
- [WY73] C. K. Wong and P. C. Yue. “A Majorization Theorem for the Number of Distinct Outcomes in  $n$  Independent Trials”. *Discrete Mathematics* 6.4 (Jan. 1, 1973), pp. 391–398. DOI: [10.1016/0012-365X\(73\)90070-8](https://doi.org/10.1016/0012-365X(73)90070-8).
- [WH19] Mischa P. Woods and Michał Horodecki. “The Resource Theoretic Paradigm of Quantum Thermodynamics with Control” (Dec. 11, 2019). arXiv: [1912.05562](https://arxiv.org/abs/1912.05562) [[math-ph](#), [physics:quant-ph](#)].



- [Woo01] William K. Wootters. “Entanglement of Formation and Concurrence”. *Quantum Info. Comput.* 1.1 (Jan. 2001), pp. 27–44.
- [YC15] Zhen-Hang Yang and Yu-Ming Chu. “On Approximating Mills Ratio”. *Journal of Inequalities and Applications* 2015.1 (Dec. 2015), p. 273. DOI: [10.1186/s13660-015-0792-3](https://doi.org/10.1186/s13660-015-0792-3).
- [YG19] Xiao-Dong Yu and Otfried Gühne. “Detecting Coherence via Spectrum Estimation”. *Physical Review A* 99.6 (June 10, 2019), p. 062310. DOI: [10.1103/PhysRevA.99.062310](https://doi.org/10.1103/PhysRevA.99.062310). arXiv: [1808.08884](https://arxiv.org/abs/1808.08884).
- [Zha07a] Z. Zhang. “Estimating Mutual Information Via Kolmogorov Distance”. *IEEE Transactions on Information Theory* 53.9 (Sept. 2007), pp. 3280–3282. DOI: [10.1109/TIT.2007.903122](https://doi.org/10.1109/TIT.2007.903122).
- [Zha07b] Zhengmin Zhang. “Uniform Estimates on the Tsallis Entropies”. *Letters in Mathematical Physics* 80.2 (May 10, 2007), pp. 171–181. DOI: [10.1007/s11005-007-0155-1](https://doi.org/10.1007/s11005-007-0155-1).



# Appendix A

## Mathematica code

The following is Mathematica code to perform the symbolic manipulations used in the proof of Proposition 5.1.6. This code was run on Mathematica version 12.0.0.0.

```
In[1]:= g[x1_,x2_,x3_,x4_] := Log[E^(x1 + 2*x2 + 3*x3 + 4*x4)
+ E^(x4 + 2*x3 + 3*x2 + 4*x1)]

In[2]:= f[x1_,x2_,x3_,x4_] := -1*(g @@@ Permutations[{x1,x2,x3,x4}]
//Total )/Factorial[4]

In[3]:= f[x1,x2,x3,x4]

Out[3]= 
$$\frac{1}{24} (-2 \log[e^{x_1+3 x_2+4 x_3+2 x_4}+e^{4 x_1+2 x_2+x_3+3 x_4}]-2 \log[e^{3 x_1+x_2+4 x_3+2 x_4}+e^{2 x_1+4 x_2+x_3+3 x_4}]$$


$$-2 \log[e^{x_1+4 x_2+3 x_3+2 x_4}+e^{4 x_1+x_2+2 x_3+3 x_4}]-2 \log[e^{4 x_1+x_2+3 x_3+2 x_4}+e^{x_1+4 x_2+2 x_3+3 x_4}]$$


$$-2 \log[e^{3 x_1+4 x_2+x_3+2 x_4}+e^{2 x_1+x_2+4 x_3+3 x_4}]-2 \log[e^{4 x_1+3 x_2+x_3+2 x_4}+e^{x_1+2 x_2+4 x_3+3 x_4}]$$


$$-2 \log[e^{2 x_1+3 x_2+4 x_3+x_4}+e^{3 x_1+2 x_2+x_3+4 x_4}]-2 \log[e^{3 x_1+2 x_2+4 x_3+x_4}+e^{2 x_1+3 x_2+x_3+4 x_4}]$$


$$-2 \log[e^{2 x_1+4 x_2+3 x_3+x_4}+e^{3 x_1+x_2+2 x_3+4 x_4}]-2 \log[e^{4 x_1+2 x_2+3 x_3+x_4}+e^{x_1+3 x_2+2 x_3+4 x_4}]$$


$$-2 \log[e^{3 x_1+4 x_2+2 x_3+x_4}+e^{2 x_1+x_2+3 x_3+4 x_4}]-2 \log[e^{4 x_1+3 x_2+2 x_3+x_4}+e^{x_1+2 x_2+3 x_3+4 x_4}])$$


In[4]:= substitute = {x1 → 5/10, x2 → 3/10, x3 → 2/10, x4 → 0}

Out[4]= {x1 →  $\frac{1}{2}$ , x2 →  $\frac{3}{10}$ , x3 →  $\frac{1}{5}$ , x4 → 0}

In[5]:= (D[f[x1,x2,x3,x4], x2, x4] + D[f[x1,x2,x3,x4], x3, x1]
-D[f[x1,x2,x3,x4], x3, x4] -D[f[x1,x2,x3,x4], x2, x1] )
/. substitute // Simplify

Out[5]= 
$$-((( -1 + e^{1/5})^2 e^{1/5} (3 - 4 e^{1/5} + 13 e^{2/5} - 13 e^{3/5} + 36 e^{4/5} - 32 e + 84 e^{6/5} - 82$$


$$e^{7/5} + 157 e^{8/5} - 127 e^{9/5} + 235 e^2 - 202 e^{11/5} + 382 e^{12/5} - 290 e^{13/5} + 500 e^{14/5} - 382$$

```

$$\frac{e^3 + 685 e^{16/5} - 513 e^{17/5} + 817 e^{18/5} - 574 e^{19/5} + 970 e^4 - 680 e^{21/5} + 1052 e^{22/5} - 691 e^{23/5} + 1114 e^{24/5} - 756 e^5 + 1114 e^{26/5} - 691 e^{27/5} + 1052 e^{28/5} - 680 e^{29/5} + 970 e^6 - 574 e^{31/5} + 817 e^{32/5} - 513 e^{33/5} + 685 e^{34/5} - 382 e^7 + 500 e^{36/5} - 290 e^{37/5} + 382 e^{38/5} - 202 e^{39/5} + 235 e^8 - 127 e^{41/5} + 157 e^{42/5} - 82 e^{43/5} + 84 e^{44/5} - 32 e^9 + 36 e^{46/5} - 13 e^{47/5} + 13 e^{48/5} - 4 e^{49/5} + 3 e^{10}}{(3(1+e^{1/5})^2 (1+e^{2/5})^2 (1-e^{1/5}+e^{2/5})^2 (1+e^{4/5})^2 (1-e^{2/5}+e^{4/5})^2 (1-e^{1/5}+e^{2/5}-e^{3/5}+e^{4/5}e-e+e^{6/5})^2 (1+e^{8/5})^2))}$$

In[6]:= N[%5, 20]

Out[6]= -0.15621966760667033050